CAPTURING COMPLEXITY: EXPERIMENTAL SYSTEMS AND EPISTEMIC SCAFFOLDS IN ANIMAL BEHAVIOR GENETICS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Nicole Christine Nelson
May 2011

CAPTURING COMPLEXITY: EXPERIMENTAL SYSTEMS AND EPISTEMIC SCAFFOLDS IN ANIMAL BEHAVIOR GENETICS

Nicole Christine Nelson, Ph. D.

Cornell University 2011

This dissertation examines knowledge production practices in the field of animal behavior genetics. Drawing on ethnographic fieldwork at a laboratory at "Western University" that uses rodents to study the genetics of alcoholism and anxiety, I investigate how practitioners establish experimental systems to model human behavioral disorders in the laboratory and manage the excess of uncertainty that they associate with the "complexity" of behavior.

To illuminate the dynamics of knowledge production in animal behavior genetics, I develop the metaphor of an "epistemic scaffold" to describe how practitioners establish and act on the conceptual foundations of particular models or tests. The image of a scaffold highlights two different processes taking place in the research community: the process of making specific links between the animal and the human using available data and theory, and the process of making more or less general claims about the utility of animal models and the applicability of animal behavior genetics findings. Methodological discussions about tests such as the elevated plus maze demonstrate how researchers negotiate about what counts as sound evidence for the connection of this test to human anxiety, and about whether researchers should claim that test models "anxiety" or only "anxiety-like behavior."

The assumption that human behavioral disorders are likely to be "complex" animates many of these discussions about the practices and conceptual foundations of the field. I analyze how researchers stabilize particular representations of multi-faceted human behaviors such as binge drinking by developing new models, and show how some researchers use these models to highlight the role of environmental factors in behavior rather than solely "reducing" human behavior to genes. Different understandings of the "complexity"

of human behavior are also associated with different expectations about the stability of animal behavior genetics experimental systems and how quickly knowledge will accumulate in the field. I demonstrate how practitioners attempt to manage expectations about what associations can be made between genes and behavior not only in the laboratory but also with other audiences in mind, such as funding agencies, policy makers, and the public.

BIOGRAPHICAL SKETCH

Nicole C. Nelson was born and raised in Waterloo, Ontario. Nicole was accepted to University of Western Ontario with a President's Entrance scholarship in 1999. Unable to decide between the sciences or the humanities, she pursued an interdisciplinary degree in Genetics and Social and Political Theory through the Scholar's Electives program. After completing her bachelor's degree in 2004, Nicole enrolled in the Ph. D. program in the Department of Science and Technology Studies at Cornell University to pursue her interests in the social studies of genetics. She received the Jasanoff Prize for the best graduate student paper in the department in 2009 for a paper related to her research on animal behavior genetics. In addition to her interest in practices in contemporary biomedicine, she has also conducted research and teaching on the intersections of science and gender and science and law. In 2008, Nicole co-organized a workshop on anticipatory knowledge in the life sciences, and co-edited a special issue of *Science and Public Policy* on anticipatory knowledge and the state arising from the workshop. She received a teaching award from the Knight Institute for Writing at Cornell for her freshman writing seminar on science, sex, and gender. Following the completion of her Ph. D., Nicole will begin a position as a postdoctoral researcher in the Department of Social Studies of Medicine at McGill University, where she will undertake an ethnographic study of cancer genomics research.

For Geoff and Judy

ACKNOWLEDGEMENTS

First and foremost, I am incredibly grateful to my academic advisors. I could not have asked for a better committee chair than Stephen Hilgartner. Throughout the process of researching and writing Steve was generous with his time and thoughts, and his visible enthusiasm for my project always perked me back up in moments when my own enthusiasm was flagging. His meticulous attention to word choice and his continual reminders to think of social processes in terms of "dynamic tensions" will no doubt stay with me for the rest of my academic career. One of the things that I will miss most about living in Ithaca is being able to head up to Steve's office for a two hour marathon meeting about my work. Michael Lynch has been a mentor and a friend from my first weeks in graduate school, and I thank him for helping me settle into life as a graduate student and for challenging me to find my own voice as a scholar. Mike's ability to write and discuss academic work in a way that is deeply insightful while still being witty (and even a little bit mischievous) is something that I admire and aspire to. Christine Leuenberger has been a role model for me as a future academic because of her talent and passion for doing ethnographic research and because of her gift for translating even dense social theory into accessible and popular undergraduate courses. Craig Simpson, one of my undergraduate professors, also deserves mention, because without his encouragement and guidance I might have ended up in a very different line of study. He was supportive (and never judgmental) as I tried on different possible career goals as an undergrad, and encouraged me to be ambitious once I found one that fit.

I am also deeply grateful to "Dr. Smith" and all of the researchers that I interacted with

during my fieldwork. My first attempts at doing ethnography were undoubtedly somewhat clumsy, and I am thankful to everyone at "Western University" for bearing with me while I figured out how to formulate my project, what it meant to do participant observation, how to ask good interview questions, and how to scruff a mouse. It was only as I was contacting other laboratories near the end of my stay at Western that I realized how fortunate I had been in finding a group of people who answered my unsolicited inquires, were willing to let me ask questions about virtually whatever I wanted, and were also interested enough in my project to make their own suggestions about "sociological moments" that I might want to observe. Their thoughtfulness about methods in their laboratory work has also made me more attentive to methodological questions in my own research practice.

Several sources of funding support made this project possible. Small grants from the Institute for the Social Sciences at Cornell and a grant on "Emerging Technologies" from the National Science Foundation (Award No. 0352000) provided me with funds do preliminary interviews that helped me conceptualize the project and obtain access to my field site. A Doctoral Dissertation Improvement Grant from the National Science Foundation (Award No. SES 0749635) did much more than "improve" my dissertation; it made it possible for me to undertake an extended period of fieldwork and travel to gather the information that I needed. I also received a Doctoral Research Fellowship from the Social Sciences and Humanities Research Council of Canada (Award No. 752-07-0739) that supported me while I was conducting fieldwork, and a Sage Fellowship from Cornell that supported me while I was writing.

I had a huge (and hugely supportive) network of people in Ithaca that helped keep me alive and well during the dissertation process, from dear friends that I made in the S&TS program to the *Gimme! Coffee* barista who made my large Americano every morning during my final writing push. Katie Proctor and Anna Geltzer, my graduate student cohort, were an invaluable source of support to me as we taught our first classes, took our exams,

and wrote our first dissertation chapters together. At one point Katie and I received a "peer collaboration" grant to observe and offer feedback on courses that we were teaching, and we were tickled at the idea of getting paid to provide the support that we'd already come to depend on from each other. My fellow S&TS students and friends Angie Boyce, Jofish Kaye, Lisa Onaga, Hannah Rogers, Janet Vertesi, Ben Wang and Emma Zuroski all provided advice, sympathetic ears and distractions at various moments in my graduate career. Hans Meyer, a fellow student who passed away while I was a student and is still sorely missed, shared many enthusiastic conversations with me about teaching and our respective projects. Members of the STS writing group (including Matt Bernius, Alima Bissenova, and Marcie Middlebrooks from the anthropology department) read several draft chapters of this dissertation, and gave me the motivation to get more words on (virtual) paper. I also thank Minnie Jardine for her weekly yoga class at Cornell that kept me sane and focused during my last years of graduate school.

My family was a key part of this support network. My brother Dan Nelson's no-nonsense advice helped me quit my worrying and get on with my work in my first years of graduate school. My sister Laura Gittens's fearlessness in the face of daunting career challenges has inspired me to be more brave in the goals I set for myself. The contributions of Harald Kliems, my wonderful partner whom I met during graduate school, were crucial in making this dissertation happen. During the final months of dissertating, he took care of literally everything else that needed doing (from laundry to references) so that I could focus on writing. I cannot thank him enough for continuing to be patient and supportive as my dissertation took over both of our lives, and I hope I can return the favor some day. And finally, my parents, Geoff and Judy Nelson, have been there to celebrate and commiserate with me through all the ups and downs of graduate school. I think they are as excited as I am to see me graduate. This dissertation is dedicated to them.

Table of Contents

List of Figures

Introduction

In June 1999, *Science* magazine published the results of an unusual study on mouse models for human behaviors (Crabbe, Wahlsten, & Dudek, 1999). Three research groups tested eight different strains of genetically identical mice in six different behavioral measures, such as activity, anxiety, and preference for alcohol. The entire set of experiments, from breeding the mice to each of the six test protocols, were conducted in parallel in the three laboratories. The authors reported that they went to "extraordinary lengths" to equate as many features of the sites as possible to ensure that the tests were performed in the same fashion (p. 1670). The mice used in the experiment were born on the same day, weaned at the same age, fed the same diet, and they slept in the same brand of bedding. Experimental equipment was built in one shop in Canada and shipped to the other laboratory sites, and the researchers negotiated detailed written protocols for how to carry out each experiment. They took special care to standardize variables that the animal behavior genetics community suspected might cause differences in mouse behavior, such as whether the mice were bred locally or shipped to the laboratory. The Edmonton laboratory that bred the mice used for the experiments shipped mice to themselves as well as to the other researchers, sending their mice on a round-trip flight to Toronto and back prior to testing. Even basic supplies, such as sheets of sandpaper used for one set of experiments, were shared between laboratories to ensure uniformity, "much to the amusement of the office staff in Edmonton who had never seen four sheets of sandpaper delivered by courier" (Wahlsten et al., 2003, p. 288).

The results of this comparison between the three laboratories varied from test to test. Some of the behavioral tests gave consistent results between the laboratories, and confirmed previous findings in the behavior genetics literature. The results of the alcohol preference experiment, where researchers measure how much a mouse drinks from a bottle of alcohol versus a bottle of water, were similar in all three laboratories, and the strains of mice that were known to have a taste for alcohol were heavy drinkers at all three sites (Crabbe et al., 1999, p. 1671). Other tests showed variations between the laboratory sites that researchers attributed to human error (or maybe more appropriately: mouse error). In the "accelerating rotarod" test, the researchers measured motor coordination by placing the mice on a rotating cylinder and gradually increasing the speed of the cylinder until the mice fell into a bin of bedding material below. The researchers covered the cylinders with their couriered sheets of sandpaper to keep the surfaces uniform between the sites, but the mice quickly discovered that under this new setup they could avoid running on top of the cylinder and use their claws to grip into the edge of the sandpaper and cling to the rod as it rotated instead. The results from the rotarod experiment, deemed "essentially uninterpretable" by the researchers, were left out of the original report (Wahlsten et al., 2003, p. 297).

But some of the tests showed differences between the laboratory sites that were not easily attributable to human error. The authors reported that "despite [their] efforts to equate laboratory environments, significant, and in some cases large effects of [laboratory] site were found for nearly all variables" (Crabbe et al., 1999, p. 1670). In Edmonton, mice of all strains scored lower on a test for anxiety-like behavior called the "elevated plus maze," generating one quippy suggestion from a Canadian reporter that perhaps the mice in Canada were simply more "laid back" (Immen, 1999, p. A6). In some cases, the differences between the sites were quite pronounced. When the researchers injected cocaine into one strain of mouse, the average increase in the mouse's movement in an activity monitor

was 701 centimeters per fifteen minutes in Albany, 667 centimeters in Portland, and over 5000 centimeters in Edmonton (Crabbe et al., 1999, p. 1672). Surprisingly, variables that the researchers suspected might influence experimental results, such as whether the mice were bred locally or shipped, did not seem to affect the outcomes at all. The authors concluded by urging readers to take into account that "for behaviors with smaller genetic effects (such as those likely to characterize most effects of a gene knockout), there can be important influences of environmental conditions specific to individual laboratories, and specific behavioral effects should not be uncritically attributed to genetic manipulations such as targeted gene deletions" (Crabbe et al., 1999, p. 1672).

The results of this multi-sited study were widely discussed in the scientific literature, and the popular press.[1] Practitioners' reactions to the experiment revealed some of the theoretical and disciplinary tensions that were especially pronounced in the behavior genetics field at the time of the study's publication. One researcher that I interviewed described the study as a kind of "Rorschach test," an ambiguous stimulus that allowed those who read it to project their hopes and fears about behavior genetics research onto the study results. Reflecting back on the publication of the study several years later, the authors reported that the reactions they received from the scientific community ranged from "indignant sigh[s] that we all knew this already to hysterical outbursts that [these] findings invalidate the entire field of behavioral genetics" (Wahlsten et al., 2003, p. 305). Some behavior geneticists saw the study results as a confirmation of both their views of behavior as a "complex" entity that is influenced by many sources of environmental and genetic variation, and of their expectations that sorting out the contributions of particular genes to particular behaviors would be a long and tricky process. The authors of the study complained that scientists outside of the field tended to interpret the study results in a much more pessimistic fashion, arguing that the study showed that animal

[1] At the time of this writing, ISI Web of Science showed over 600 citations of this study in the scientific literature.

behavioral genetics was just not a reliable science and would not be likely to succeed in linking specific genes to behavioral disorders (Wahlsten et al., 2003, p. 306). And while the discussion of the study in the scientific literature focused on the methodological implications of the study's findings, the popular press and some scientists also used the study as an opportunity to discuss the public understanding of genetics. The European Molecular Biology Association, for example, used the study in a statement discussing the genetic basis of intelligence to argue that

> the waters of reality are muddier than the media would wish the public to believe … despite the fact that we can now identify more predisposition genes than ever before, we must resign ourselves to the fact that they will never be the whole story. (Moore, 2000, p. 101)

Other articles in the news media characterized the results as "one for nurture" in an ongoing contest between the forces of genes and the forces of environments. As one newspaper headline put it succinctly, "mice study shows genes are not always destiny" (Reuters News, 1999).

This study was also my entree into the world of animal behavior genetics. I had discovered it after reading an article about the effect of the laboratory environment in animal behavior genetics in *Discover* magazine (Yeoman, 2003), and I too pondered the study's meaning for understanding scientific practice. What had motivated these researchers to undertake an experiment that would reveal some of the uncertainties and contingencies in their own knowledge production processes? Were they attempting to make a case for greater standardization of testing practices or more attention to the laboratory environment? Or were they arguing for a more substantial reformulation of the way that researchers conceptualized gene action by focusing on the interaction of genes and environments? And why did they not seem to be concerned about whether this study would damage the credibility and authority of their field?

The goal of this dissertation is to investigate how animal behavior geneticists use animal models, especially mouse models, to produce knowledge about "complex" human behavioral disorders such as anxiety and alcoholism. The laboratory mouse is the most widely used model organism in biomedical research today, so widely used that Creager, Lunbeck and Wise argue that "at the dawn of the twenty-first century, the face of biology may well be that of a laboratory mouse" (2007, p. 1). But as the multi-sited experiment I described demonstrates, using the mouse to generate knowledge about human disorders, especially disorders that are considered to be "complex disorders" by researchers, is not a straightforward task. Mice offer many opportunities for studying the genetics of disorders like alcoholism because of the vast array of genetic resources associated with the mouse, but they also present many challenges. It is difficult to get mice to consume alcohol voluntarily, to standardize housing conditions between laboratories even on the same campus, to get repeatable results between laboratories, and to produce tests that share some resemblance with human symptoms. Practitioners also have different expectations about what kinds of associations they will be able to make between genes and behavioral disorders, what kinds of methods and practices will be needed to produce stable knowledge, and how quickly knowledge will accumulate in the field.

With these issues in mind, I investigate the following questions in this dissertation: What do animal behavior geneticists mean when they talk about behavioral disorders as "complex" disorders? How do expectations about the complexity of behavior shape their research practices, the way that they socialize new members into the field, and their expectations of what kind of knowledge they will be able to produce? Exploring how researchers describe the challenges posed by the "complexity" of behavioral disorders helps illuminate both their understandings of good research practice and of gene action. I also explore how animal researchers attempt to convincingly connect the knowledge that they make in the laboratory using animals to human disorders. When do they make strong

claims about the relationship between their animal models and human disorders, and when do they insert distance or uncertainty into this relationship? What kind of vision of the human comes out of animal model research, and does it "reduce" the human to genes? Finally, I am also interested in investigating how animal behavior geneticists' knowledge communities and practices are shaped by other cultures and practices. How do researchers stabilize their facts and talk about their work with respect to different audiences, such as the research community, funding agencies, and the public? How do researchers conceptualize the political cultures that they work in and how do they participate in public discussions about the meaning of genetics?

By examining the knowledge production practices of animal behavior geneticists, I aim to contribute to ongoing discussions, both in the field of science and technology studies and elsewhere, about the form and consequences of contemporary genomic research. Much has been written about the ethical, legal, and social implications of human genetics research, and behavior genetics research in particular (see, for example, Lewontin, Rose, & Kamin, 1984; Duster, 1990; Parens, Chapman, & Press, 2006). Commentators have expressed divergent views on how to understand contemporary behavior genetics research and its future applications. Peter Conrad (1999a), for example, likens the genetic explanations of the late twentieth century to the explanations of health and illness offered by germ theory in the early twentieth century. He argues that like germ theory, genetic discourse promotes the flawed idea that there is a single, biological causal agent for behavioral disorders. Scholars have also argued that the tendency of behavior genetics research to "reduce" humans to genes could have detrimental effects on public policy and popular thinking about genes. Commentators argue that genetic research ignores important environmental factors that shape behavior in its quest to find biological and genetic determinants, and could promote questionable social polices or applications (Lewontin et al., 1984; Duster, 1990; S. Rose, 1997). Nikolas Rose (2007), on the other hand, argues that assumptions

about gene action in contemporary genomics are quite different from early twentieth century research on heritability. He points out that contemporary genetic research is based on principles of susceptibility and probability, not a straightforward genetic determinism, but he argues that the way that genetic research is applied nonetheless deserves critical scrutiny.

Studying the knowledge production practices of animal behavior geneticists ethnographically offers a different way of analyzing the assumptions about genes, environments, and human behaviors that are embedded in behavior genetics research. In this dissertation, I aim to describe how animal behavior geneticists conceptualize gene action, frame the human by selecting particular features of human behaviors to focus on, and envision the future uses of behavior genetics facts. Rather than evaluating the merits of the claims that animal behavior geneticists make or speculating about the future consequences of animal behavior genetics research, I explore how practitioners themselves negotiate a contested terrain of expectations and understandings of what the genetics of behavior looks like in their day-to-day work in the laboratory. In this sense, this dissertation is less about the manufacture of specific facts than it is about the *management* of the perceived "complexity" of the genetics of behavior.

One area that I devote particular attention to is how animal behavior genetics researchers manage the relationship between the animal and the human in their work. In science and technology studies, scholars have explored how particular organisms came to be used as "model" organisms (Ankeny, 1997; Creager, 2002; Rader, 2004; Leonelli, 2007), and how knowledge communities that use particular model organisms function (Kohler, 1994; Shostak, 2007). Rader (2004), for example, details how genetically standardized inbred strains of mice came to be widely used as biomedical tools first for studying cancerous tumors and eventually as models for studying human biology more generally. Shostak (2007) explores how different knowledge communities interact around the mouse. She

argues that as a widely used biomedical tool that can be used in many different ways, the mouse allows researchers from widely divergent disciplines to interact with each other and generate shared standards and practices. Scholars studying animal models have also offered different ways for thinking about the way that animal models model, as a process of "extrapolation" (Schaffner, 2001), providing "exemplars" (Creager, 2002), or through a process of "case based reasoning" (Ankeny, 2007).

This literature on animals as "models" in the biomedical sciences offers resources for thinking about how animal behavior geneticists link the animal and the human together in their laboratory work. By asking how mice in mazes "model" human experiences such as anxiety disorders, I aim to examine how practitioners make and manage relationships of similarity between their animal experiments and human behaviors. I explore how researchers link the mouse and the human together in specific ways to provide justification for their research programs and also emphasize differences between their models and human behaviors. I also investigate the representation of the human that is embedded in animal behavior genetics research programs, and how this representation changes as researchers build and modify their experimental systems.

A final contribution of this dissertation is to characterize animal behavior genetics as a system of knowledge production with respect to different communities of practice, institutions, and sociopolitical contexts. One criticism of some studies of knowledge production in science and technology studies is that they do not take into account how institutional factors and cultural contexts also shape the way that knowledge is produced in laboratories. Kleinman (1998, 2003), for example, argues that laboratory ethnographies have tended to employ microsociological approaches that focus on the agency of scientists as producers of knowledge and ignore institutional and structural features outside the walls of the laboratory that shape the way that scientists act. In his study of a plant science laboratory, Kleinman uses ethnographic methods to explore how macrosociological

trends outside of the laboratory, such as the increasing commercialization of biological research, shapes practices inside of the laboratory. Tuunainen (2001) likewise argues that Rheinberger's (1997) description of how scientific knowledge is produced through work with "experimental systems" is too internalistic. Using his study of a potato biotechnology research group, he argues that research programs are directed not only by epistemic concerns but also by practical considerations and concerns about the societal relevance of the research, and that these factors should also be considered part of the experimental system.

Mouse models are also developed and stabilized not only in laboratories but with other communities and audiences in mind, such as the broader behavioral genetics and neuroscience communities, funding agencies, and the larger social and institutional worlds in which behavior genetics work takes place. With these critiques in mind, I pay attention to the wide variety of policies, actors, and institutional structures that act on the body of the mouse in the arena of the laboratory, and I examine the ways in which animal behavior geneticists themselves talk about other researchers, "the public," and the sociopolitical context in which they work.

Background

There are several reasons why animal behavior genetics is an especially appropriate place to investigate questions about how mice are used to produce biomedical knowledge about the human. First, as the reaction to the multi-sited study shows, there are active discussions within the animal behavior genetics community about what kinds of facts practitioners can reasonably expect to produce using animal models and how quickly and easily they will be able to do so. When the study was published in 1999, it touched on anxieties about methods and the meaning of behavior genetics results that had been building in the field for some time. In the early 1990s, the introduction of new techniques for investigating the mouse genome led to a dramatic expansion of the number of laboratories studying genes

and behavior in mice. Using methods for "knocking out" genes in the mouse genome, researchers could make mice who had a non-functioning copy of a specific gene (known as the "knockouts") and compare them with mice who still had a functioning copy of the gene (known as the "wild type"). Although similar strategies for comparing mutants to wild type animals had been used in genetic research programs for years, the knockout technique allowed researchers to target *specific* genes, rather than causing mutations at random places in the genome or waiting for mutations to occur spontaneously.[2] Knockouts offered a powerful way to experimentally demonstrate the relationship between genes and behavior, and within a short period results from knockout studies flooded the behavioral literature.[3] Better maps of the mouse genome and improved statistical techniques also provided researchers with new options for searching for genes associated with particular behaviors. Quantitative trait loci mapping (QTL) techniques, which used genetic maps and statistical models to identify regions of the genome where there might be a gene that influenced a particular trait, were also adopted by animal behavior geneticists in the early 1990s.

Researchers saw these methods as full of promise for finding genes associated with particular behaviors.[4] One animal behavior geneticist I interviewed described the combined effects of these new techniques as nothing short of "revolutionary." She said:

> Twenty five years ago, all we could say was "oh yeah, there's a genetic component, and we can do these complicated breeding schemes, and probably eight genes control this trait, but we don't really know if we're confident about that estimate, and we have no idea what it is or where it is or how to find it," you

[2] Researchers using other model organisms such as *Drosophila* and *C elegans* built highly successful genetic research programs by using spontaneous or randomly generated mutants to map the genomes of these organisms and associate particular genes variants with physical characteristics.

[3] A lab at MIT headed by molecular biologist Susumu Tonegawa published the first behavioral knockout study in 1992 showing that knocking out a gene for an enzyme found in the synapses of the hippocampus affected the mouse's ability to learn and remember tasks, and the Tonegawa lab alone went on to publish seven more papers on the behaviors of knockout mice in high profile journals, such as *Cell* and *Science*, in the following years (A. Silva, Paylor, Wehner, & Tonegawa, 1992).

[4] For a review article published in the mid-1990s outlining some of the possibilities for how QTL, knockout, and other related techniques might be used in the alcohol research field specifically, see Crabbe, Belknap, & Buck, 1994.

> know. And all of this has happened, this actual gene finding, in the course of my existence in the field.

Rather than just describing disorders in terms of genetic and environmental contributions, scientists could now engage in "actual gene finding," looking for regions of chromosomes and perhaps even specific genes that were linked to behaviors or disorders. QTL mapping helped researchers find potential regions (and hopefully potential genes) for the behavior that they were interested in, and knockout studies offered a way to experimentally test the relationship between a gene and behavior by demonstrating what would happen if the gene was rendered non-functional.

But at the same time that researchers optimistically explored the possibilities offered by these new techniques, there was also uncertainty in the field about what these methods could produce. Some of the findings that initially caused so much excitement in the animal behavior genetics community, such as the results of early knockout studies, couldn't be replicated when the mice were sent to other laboratories for study. In one laboratory that I will explore in more detail later, researchers found that they could not replicate the results of their own study with a knockout mouse when they did the same experiments only a few years later. One senior behavior geneticist recalled the situation as follows:

> What we started seeing in the early years of the genetic engineering revolution with mice was study after study where they would look at aggression using a simple task. And what we started to see was that pretty much every gene that was knocked out could be called an aggression gene. But then there were some notable failures to replicate, where the same knockout, when sent to somebody else's lab, they would find something different.

While many practitioners assumed that the differences were probably due to idiosyncrasies in the ways that individual laboratories performed their experiments, the multi-sited study showed that even incredibly carefully executed studies conducted by well-respected laboratories could still produce different results. Debates around how to interpret the results of knockout studies in the 1990s and ongoing discussions around other experimental techniques to associate particular genes with behaviors offers opportunities for exploring

researchers' understandings and expectations about what kinds of associations they think they will be able to make between genes and behaviors and what techniques they think will be needed to do so.

The field of animal behavior genetics also provides a good site for studying how researchers using animal models make connections between their research with animals and human disorders. As in many areas of biomedical research, animal behavior geneticists use the mouse as a tool for understanding human behavior, neurobiology, and genetics. Establishing a plausible relationship between mice and humans is especially challenging in animal behavior genetics, where researchers are using animals to model disorders that even they describe as "uniquely human" (see, for example, Powell & Miyakawa, 2006). As Gusterson (2008) might put it, there appears to be a "surplus of ambiguity" (p. 559) around the connection of the mouse and the human in contemporary animal behavior genetics. In his study of nuclear weapons science, Gusterson argues that the ban on testing nuclear weapons creates a perceived "lack" in the weapons science community that poses a problem for knowledge production. Without access to test data—the kind of data that practitioners believe is necessary to answer particular scientific questions—there is fundamental disagreement on key technical issues in the research community, such as whether existing warheads or new designs will be more reliable. Likewise in animal behavior genetics, practitioners describe the task of studying the genetics of human behavioral disorders as quite difficult because they lack access to key kinds of information. Animal behavior geneticists argue that developing animal models is difficult because the human disorders themselves are heterogeneous and ill-defined, and because they cannot access information about the subjective experience of animals, a key component of how behavioral disorders are measured in humans. At the same time, researchers describe animal models as absolutely necessary for understanding behavioral disorders because of a lack of tools to adequately study the human brain. Examining changes in gene transcription or in

neurotransmitter levels in humans is also difficult because researchers cannot easily access brain tissue. The tensions that practitioners describe around the relationship of animal models to human behavioral disorders makes the field of animal behavior genetics a good place to study how researchers negotiate the relationship of the animal to the human and use model organisms to produce biomedical knowledge.

Finally, the field of behavior genetics also offers opportunities to study the relationship of technical practices with their sociopolitical context. The coverage of the multi-sited study in the popular press demonstrates both the social salience of behavior genetics results and how scientists themselves use behavior genetics information to contribute to public discussions about the meaning of genetics. Neuroscientist Robert Sapolsky (2005), for example, used the multi-sited experiment as an example in a popular essay on "gene hyping" to illustrate some of the complexities of the relationship between genes and behavior. After carefully explaining the experimental setup and outcomes, he concludes that:

> The moral is that one should not get too excited about some new genetic component of behavior until the effect has been replicated in a number of places and with a broad array of tests—something that is seldom done. … The conclusion must be that many published accounts linking groups of genes to specific behaviors could be well off base. Don't get me wrong and overestimate how much I'm trying to bash genes … but amid our current near-feverish interest in genes, especially among the lay public, it's worth noting that the emperor is a bit less accessorized than usually assumed. The environment, even a subtle one, can still more than hold its own in the biological interactions that shape who we are. (Sapolsky, 2005, p. 36–37)

Sapolsky's interpretation of the meaning of the multi-sited study is very much in line with the conclusions of the study's authors, although his message is directed at a different audience and intervenes in a different debate. Like Sapolsky, many behavior geneticists also believe that the public ascribes too much power to genes. The field of behavior genetics has been the subject of several highly publicized and politicized controversies about its methods and findings, such as debates about race and intelligence and genetic links to

sexual orientation, and behavior geneticists are particularly attuned to issues about the public perception of their field and potential uses of their research in public policy. The social salience of behavior genetics knowledge and the awareness and participation of practitioners in public discussions about the meaning of genes makes behavior genetics a good site to study how cultural context impacts research practice.

The Smith Laboratory

To better understand how researchers generate knowledge about the genetics of human behavioral disorders using animal models, I spent a semester as a participant observer in a drug and alcohol genetics lab in a university that I will refer to in this dissertation as "Western University."[5] Western is situated in a hilly city on the West Coast of the United States, and the laboratories in the Department of Neuroscience are spread out in many buildings across campus and in the buildings of the affiliated teaching hospital. During my time at Western University, I was hosted by the Smith laboratory, a relatively small lab with many more mice than people. At the time that I was there, Dr. Smith employed about half a dozen researchers and technicians, was training several students and postdocs, and his team was breeding or housing thousands of mice. The laboratory is located on the same floor as many of the rooms where the mice are housed at Western, and the first impression I had of the laboratory was the distinct (and not especially pleasant) smell of mouse that greeted me as I stepped off the elevator.

Behind a locked security door, the Smith laboratory's work spaces are nested in between some of the most highly trafficked areas in the building. The meeting and lunchroom where many of the laboratory employees have their desks is across the hallway from the cage washing room, where the animal care staff bring hundreds of mouse cages to be

[5] I use pseudonyms for the university and the researchers who work there throughout this dissertation to protect the anonymity of those I studied. See appendix A for more information about interviews, names, and pseudonyms.

sterilized each day. Next door to the lunchroom are several "colony rooms" housing thousands of mice in labeled cages, and around the corner are the "procedure rooms" that hold special equipment such as fume hoods and activity monitors and provide space for doing behavioral experiments. On any given day, the hallways outside of the Smith laboratory rooms are filled with technicians and graduate students from many different laboratories, all dressed in green scrubs, wheeling cages of mice draped in sheets off to procedure rooms for experiments or wheeling clean cages with fresh bedding into the colony rooms.

Dennis Smith has worked for more than thirty years in the behavior genetics field, and nearly all of his career with mice. Although he spends most of his time working at his desk and rarely touches a mouse now, he talks affectionately about his research subjects, hopping out of his office chair to simulate a jumping mouse or joking about strains of mice that are so inactive that they could barely muster the energy to run away from you. Dennis is down-to-earth, funny, and generous with his time—characteristics that make him an excellent scientific colleague as well as an excellent host for a curious sociologist. Dennis enjoys close working relationships with many of the other professors at Western; and mice, money, ideas, and people flow relatively freely between the Smith laboratory and other collaborating laboratories. Laura Martin, another professor at Western whose laboratory is in the same building, is a longtime collaborator of Dennis and the two have a tradition of sharing lunches together to talk about their work and brainstorm new projects when they are both in town. While there are some notable differences in the content of work done in the Smith laboratory and other laboratories, I often talk about the Smith laboratory and its collaborators at Western nearly interchangeably in the dissertation because they are so closely related, especially in terms of the way that they think about methods.

The Smith laboratory and its collaborators use mouse models to study the genetics and

neurobiology of addiction disorders, as well as related behavioral disorders or traits, such as anxiety and impulsivity. The aim of the research conducted at Western is to identify regions of the mouse genome or particular genes that make mice sensitive to particular drugs or susceptible to addiction, and if any specific genes can be identified, to study the "mechanisms" of how these genes make animals more or less susceptible to addiction. Even though scientists at Western describe their work as "basic science," it has a clinical orientation. Ultimately, researchers in animal behavior genetics hope to generate information that will eventually be useful in studying human behavioral disorders, although there is some disagreement about what kind of information is most likely to translate into clinical uses and how long this translation will take. The Smith laboratory and its collaborators use many different methodological techniques to try to make associations between genes and behaviors, from techniques that one of Dr. Smith's technicians described as "old school," such as selective breeding (where researchers mate mice with high or low scores on a particular test to create mice with particular behavioral tendencies), to techniques requiring expensive and high-tech equipment such as microdialysis (where researchers use a small tube inserted into the mouse's brain to measure neurotransmitter levels as the mouse is doing a behavioral test). While some labs at Western focus primarily on molecular biology techniques, many researchers at Western describe themselves as "behaviorists" because they primarily study live, "behaving" animals first and then look for whether or not genetic differences influence those behaviors. Researchers talk about the skill set needed to successfully design and perform behavioral experiments as something that is quite different from doing "bench" techniques that take place in a wet lab. Dennis in particular is known for designing studies involving a hundred or more mice at a time, and executing these studies requires the careful coordination of several technicians as well as the management of vast numbers of tipsy mice.

The Smith laboratory researchers' identification as "behaviorists" is important for un-

derstanding their location in the varied topography of the research field. Panofsky (2006) describes the contemporary behavior genetics field as an "archipelago" of related specialties clustered together under the banner of "behavior genetics," but having distinct training, methods, and professional affiliations. He identifies four different "islands" that comprise the behavior genetics field: psychological behavior genetics, psychiatric genetics, molecular genetics and animal behavior genetics. Panofksy describes animal behavior genetics as the most heterogenous in terms of training, professional position, and research practices, with some researchers identifying primarily as "behavior geneticists" and others who do work that could be classified as behavior genetics but identify as molecular biologists or neuroscientists. Dennis, for example, had not been to the annual meeting of the Behavior Genetics Association in more than a decade, but said he still maintains his membership because he says he "believes in the idea of behavior genetics." The majority of the researchers that I interacted with identified themselves as behavior geneticists or behavioral neuroscientists, worked primarily or exclusively with animal models, and had backgrounds in psychology or animal behavior. These identifications are important because researchers themselves use disciplinary affiliations to align themselves in particular methodological debates.

The ethnographic and interview information that I collected at Western University forms the core of my dissertation, and informed my selection of later research sites and interviewees. I conducted an initial visit to Western University in September 2007, lasting a few days, and spent one semester as a participant observer at Western from January to May 2008. While at Western, I attempted to involve myself in as many aspects of scientific life as possible, focusing in particular on courses, training sessions, or setting up new experiments where discussions of methodological issues might be especially prominent. I attended the weekly laboratory meetings where recent experimental results were discussed and new meetings were planned, and I took detailed notes during these meetings. Dennis

graciously incorporated me into his weekly meeting schedule for his graduate students and postdocs, where we discussed events taking place in the laboratory as well as general topics such as the review process for scientific papers, scientific organizations and collaborative projects that the Smith laboratory participated in, past research projects, and future projects for the laboratory. I took the introductory behavior genetics class, co-taught by Dr. Smith and Dr. Tremblay, with the new group of graduate students in the behavioral neuroscience program. I also took the mouse handling course for new laboratory practitioners, where Aiden from the animal care staff patiently walked me through the process of injecting my very own cage of mice with saline solution. In the laboratory, I observed experiments conducted by current graduate students and technicians and sometimes I acted as the scribe for these experiments, writing down mouse weights and quantity of alcohol consumed. I also helped with some of the day-to-day tasks of the laboratory, such as filling up bottles of alcohol solution, filling test tubes in preparation for genotyping tissues, preparing mouse cages for upcoming experiments, scoring behavioral data, and proofreading data sheets. Finally, I spent a good deal of time hanging out at the lunchroom table, the "heart" of the Smith laboratory, where grad students and technicians tended to congregate before and after experiments. Throughout this process, I took field notes on the daily activity of the laboratory (often in a cafe on campus before heading home for the day), and conducted semi-structured interviews with researchers, graduate students and other members of the research community. In total I interviewed 24 individuals at Western (a full description of interview methods follows in the next section).

While I was based in the Smith laboratory for most of my time at Western, I also did "guest visits" where I observed a week's worth of activities in the laboratories of two of Dennis's close collaborators, Laura Martin and Ruth Tremblay. At the beginning of these weeks I presented myself and my research and told them the types of things that I was interested in watching (such as training sessions, troubleshooting or setting up new

experiments, or experiments in progress). Various members of the laboratory then offered to let me shadow activities that were taking place that week, and in some cases suggested activities that I might be interested in seeing that I had not thought of (such as how mouse housing rooms were run in a different building on campus). While brief, these visits were especially productive, and I was able to observe training sessions for high school students, new graduate students, and entire laboratories being trained to use new pieces of equipment.

My research in the Smith laboratory follows in the tradition of laboratory ethnographies in science studies, where sociologists and anthropologists spend time as participant observers to gain insight into how scientific practice unfolds in the day-to-day work of the laboratory (Latour, 1986; Lynch, 1985; Knorr-Cetina, 1981; Knorr-Cetina, 1999; Traweek, 1988). These studies demonstrated, amongst other things, that scientific practice is organized much differently than scientific papers might suggest, and that by studying only the scientific literature scholars interested in knowledge production might risk mistaking the literary conventions of scientific writing for what the practice of science is actually like (Lynch, 1985). By observing the everyday work of the laboratory, ethnographers showed that scientific experiments often don't go as planned or don't work at all; that theorizing, data collection, and hypothesis building often happen all at the same time; and that results of experiments are not clear and readable presentations of nature but need to be made clear by researchers. Laboratory studies offered descriptions of the processes by which scientists generate stable facts that were quite different from canonical descriptions of the scientific method. Latour (1986), for example, described the laboratory as a system for generating literary "inscriptions" that gain more credibility and solidity as they circulate in the scientific literature. Lynch (1985)described knowledge production in the neuroscience lab that he studied as a kind of craft work, where researchers sort facts from artefacts and resolve disagreements about what counts as data through local interactions and practices.

For my own study, I employed the method of studying scientific knowledge production through participant observation in a laboratory to gather information on both the challenges that practitioners identify in their day-to-day work and on the local, social aspects of knowledge production.

Additional Field Sites and Data Analysis

Following my fieldwork at Western, I visited numerous other laboratories to collect more information on particular methodological issues I was interested in pursuing. In total, I visited five additional laboratories in the United States, three laboratories in Canada, and two in Germany. Visits to these laboratories lasted on average 1-2 days. I conducted interviews with laboratory members, and asked for a tour of the laboratory and permission to observe ongoing experiments if any were taking place while I was there. The members of these laboratories that I interviewed (usually principal investigators) had all participated in some way in methodological discussions in the field, such as by publishing textbooks or highly cited papers on methodological issues. One site that I visited was a phenotyping center that performs behavioral analysis and does training for other laboratories. I also interviewed other individuals who were relevant to my topic, such as a veterinarian working for a major mouse supplier and officials at the National Institute on Alcoholism and Alcohol Abuse (NIAAA) and the National Institute on Drug Abuse (NIDA).

Although these additional laboratory visits and interviews sometimes offered informative contrasts to my fieldwork at Western, they were not intended to be explicitly comparative but selected to allow me to gather more data around core theoretical categories. In some cases I interviewed individuals whose work contrasted with that of the Smith laboratory (such as laboratories working in more explicitly "translational" settings or with different model organisms), and in other cases I selected interviewees because of their participation in shared research projects or methodological debates (such as

researchers who were members of the Alcohol Research Consortium, which I will discuss below). In many cases, researchers who had shared methodological concerns also had disciplinary backgrounds and opinions that were quite similar to those that I interacted with at Western, and as such these additional interviews do not constitute a representative sample of the field. This was abundantly clear during one interview with a Canadian researcher where I told him that I had heard quite a lot about a particular methodological issue, and he responded that was only because I was in the "echo chamber" with other people who were interested in methodology.

Two other sites where I collected data were the Mouse Phenome Project (MPP) at the Jackson Laboratory and an Alcohol Research Consortium funded by the NIAAA. The MPP is an initiative to gather "phenotype" data, that is, information on the physical and behavioral characteristics of a select group of inbred mouse strains. The MPP researchers also collect information on the protocols used to gather that data and the housing conditions that the mice are kept in. Some of the behavior geneticists I interviewed in my initial field work for this project were contributors to this database, and I selected this site to investigate how animal behavior geneticists interacted with other researchers from different disciplinary backgrounds around a shared research tool (the laboratory mouse). I spent a week at the Jackson Laboratory in January 2008 and interviewed 5 individuals while I was there. I also gathered information on an Alcohol Research Consortium group, one of several consortium projects funded by the NIAAA. It had been running for approximately eight years at the time I observed it, and brought together members from about 20 laboratories from across the United States (including members from Western University) to work on a common animal model. The backgrounds of the consortium's members were wide-ranging, from researchers who were major players in the alcohol research field to researchers who specialized in electrophysiological studies of cells in the brain. I observed one of the semi-annual meetings of this ARC in June 2008 and interviewed about a third

of the senior researchers involved in the project (9 interviews), focusing on researchers who were working primarily with mouse models.

In total, I conducted interviews with 52 unique individuals between 2006 and 2009, and some individuals were interviewed multiple times. These interviews were recorded using a digital audio recorder and conducted in person when possible. Two interviews were conducted over the phone, and five interviews were conducted in person but not recorded due to varying circumstances (such as participants who did not want to be recorded, or interviews that took place in a venue where recording was not possible). For unrecorded interviews, rough notes were recorded during the interview and detailed notes about the interview were recorded within one day of the interview (and usually immediately following the interview). The interviews lasted on average one hour, and in some cases took up to three hours. Interviews with graduate students and technicians were conducted confidentially, and principal investigators were given the option of being interviewed confidentially or non-confidentially. Appendix A provides a complete list of the people that I interviewed, the dates of interviews, a general description of their position, and the pseudonym by which they are identified in the text.

All interviews were conducted in a semi-structured, conversational fashion using an interview guide. In a few cases I attempted to ask similar questions in all interviews (for example, I asked almost all interviewees to tell me five things that they considered to be important to control for in their laboratory work), but in most cases the wording of the questions, the order of the questions, and the topics covered varied in each interview. I used the same interview guide for all interviews with graduate students, and these interviews were conducted as a rough "oral history" of their scientific careers. I began by asking them to tell me how they became interested in a career in research; about their first experiences in the laboratory; who taught them the techniques that they used and what kinds of things were emphasized to them during training; what techniques they didn't

like doing; and what techniques they found easy or difficult to learn. For students who had worked in multiple laboratories or had completed their first-year rotations through several laboratories at Western, I asked them to compare the laboratories that they worked in. I also asked graduate students about what they thought about the validity of the animal models that they were using, how they envisioned their research would be used in the future, and how they discussed their work with friends and family. For interviews with principal investigators and other actors, I developed a unique interview guide based on the participant's area of research and participation in different methodological discussions. In some cases I asked participants to tell me about their research histories, but in most cases I used particular methodological issues that were relevant to their work as the starting point for interviews. I also asked these researchers questions that were similar to those asked of graduate students about model validity, future applications, and the public communication of behavior genetics research.

In addition to these interviews and laboratory visits, I also attended three scientific conferences as an observer. The selection of these conference as sites for observation was informed by the disciplinary backgrounds of the researchers at Western and by the focus of these conferences on methods. I attended the 2008 annual meeting of the Behavior Genetics Association (BGA), the longest standing professional organization for behavior geneticists. Several researchers at Western were past or current members of the BGA, but no researchers from Western attended the conference in the year that I went. The conference provided an informative contrast to the work I had been observing at Western, since the conference was almost entirely focused on psychological approaches to studying human behavior genetics. The sole animal researcher who presented at the 2008 meeting (who was researching sex specific behaviors in *Drosophila*) noted at the beginning of his talk that he was "very disappointed" about the noticeable lack of animal researchers on the agenda. I also attended the annual meeting of the International Behavioral and

Neural Genetics Society (IBANGS), a society formed in 1996 by a group of animal behavior geneticists who felt that the BGA was no longer meeting their needs—scientifically or politically—, following the controversy caused by BGA president Glayde Whitney's presidential address on racial group differences in humans. I attended the full conference as well as the satellite session held before the conference on impulsivity. The 2008 IBANGS meeting attracted primarily researchers working with mice and rats, but human behavior geneticists had a small but visible presence. Western researchers are deeply integrated into the social networks of IBANGS, and the attendees of the conference included many current and former collaborators, students, and postdocs from Western. Finally, I also attended the 2008 Measuring Behavior conference, a biennial meeting held in the Netherlands that focuses on methods and techniques in behavioral research. This conference drew a wide audience of researchers from both North America and Europe, and from fields as diverse as ethology, behavioral neuroscience, psychology, and human-computer interaction. The conference included teaching workshops designed to introduce new practitioners to the existing tests for assessing particular behaviors, and I attended one workshop focused on mouse behavioral testing. Measuring Behavior was sponsored by a commercial manufacturer of behavioral testing equipment, and it was the only conference I attended that had a significant commercial vendor presence.

I also used published documents both to select interviewees and to deepen my analysis of particular methodological issues. I used literature searches to identify researchers who had published commentaries, review articles, or research papers touching on methodological debates that were of interest to me, such as discussions about the "background effect" or the validity of the elevated plus maze. Before each interview, I also used researchers' publication histories to prepare guides for the interview. Finally, I also used document analysis to look at how the tests that I focused on, such as the elevated plus maze, were discussed in the literature. Published documents also informed my analysis of the ARC's

development of mouse models of binge drinking, but I do not cite these documents here for confidentiality reasons.

To facilitate data analysis, I transcribed all recorded interviews in full, and entered interview transcripts and field notes into the software program ATLAS.ti (ATLAS.ti Scientific Software Development GmbH, 2010). My approach to qualitative data analysis was guided by grounded theory principles of descriptive coding, theoretical coding, and constant comparison (Glaser & Strauss, 1967). I started by coding a sample of my interview transcripts, freely generating descriptive codes that reflected the topics, terms, concepts, and problems that were present in the interviews. The goal of this stage of open coding was to develop a sense of what is important in the worlds of the researchers that I interacted with, what they talked about frequently and directed their attention to, what their main concerns and problems were, and what methods they used to solve them (with the very important caveat that the information that they provided was offered in response to my questions and directed to a large extent by my research interests). This process of open coding generated a list of about 150 codes, and from this list I developed a focused list of codes that were both "grounded" in the data and of theoretical interest to me. I then coded the remaining transcripts and field notes using this set of focused codes, refining the codes and adding new ones where necessary. I also used ATLAS to auto-code some key terms in the data. For example, I used auto-coding to identify instances where researchers used permutations of the term "complexity" or "complex" to describe some aspect of their work. Throughout this process, I kept in mind the grounded theory principle of "constant comparison," which encourages the researcher to continually compare new information and quotes that are added to a category to other information that has been coded with the same category. I looked for consistencies in these categories (such as factors to control for that were mentioned over and over), compared different instances of the same topic or technique (such as the ways that researchers control mouse genomes versus mouse

housing, the laboratory test environment, or their own bodies), and data points that were most divergent in addition to those that were most similar (such as extremely conservative or permissive attitudes towards controls).

Finally, I also employed some of the "cartographic" approaches to data analysis described by Clarke (2005). Clarke argues that one of the limitations of grounded theory is that it focuses on concerns raised by actors, but does not offer a way to investigate silences and absences in the data. She offers several techniques as supplements to grounded theory that allow researchers to investigate the "situations" in which action takes place. In particular, I employed Clarke's techniques of "positional mapping" to lay out the major arguments made or positions taken in particular debates or controversies, and "social worlds/arenas maps" that identify key groups of actors and non-human elements that interact within a particular "arena." The technique of positional mapping is useful for visualizing the similarities and differences between the arguments and positions that actors take, as well as for identifying arguments or positions that are not taken in the data set (for example, the notable absence of the argument in my data set that mouse anxiety is *the same* as human anxiety). The technique that I employ in chapter 4, following the "career" of the laboratory mouse, is similar to Clarke's notion of mapping social worlds or arenas. By following the mouse through the arena of the laboratory, all of the different people, instruments, regulations, and discourses that come into contact with the mouse become visible. This mapping of the laboratory provides a basis for asking questions about why parts of the arena of the laboratory are particularly prominent in actors' discourses (such as institutional policies for controlling disease and experimental handlers) and why some aspects are nearly absent (such as the animal care staff).

Outline of the Dissertation

I begin my discussion of the social worlds of animal behavior genetics research by introducing one of the main problems that animal behavior geneticists identify in their work in the laboratory: how to manage the "complexity" of behavioral disorders in order to produce stable knowledge about genetic contributions to particular behaviors. *Chapter 1* explores what researchers in the Smith laboratory mean when they say that the disorders they study are "complex," and how these expectations about the complexity of behavioral disorders shape their knowledge production practices. Drawing on Collins's (1985) and Rheinberger's (1997) descriptions of knowledge production in the laboratory, I argue that researchers in the Smith laboratory are developing local standards for what counts as acceptable experimental practice as well as expectations about how stabilized their experimental systems are and how quickly knowledge will accumulate in the field. I describe some of the social practices that contribute to the formation of these accepted practices and expectations, such as the process of training new practitioners to run experiments and the circulation of "cautionary tales" about past experiences in the laboratory, and some of the differences in the way that researchers talk about control of different aspects of their experimental systems.

After characterizing some of the general features of knowledge production in the Smith laboratory, in *chapter 2* I explore specifically how researchers use animal models to produce knowledge about the human. Using researchers' discussions about the validity of a particular behavioral test, the elevated plus maze, I develop the concept of *epistemic scaffolding* to describe how knowledge about the human is produced using animal models. Building on existing descriptions of how knowledge is produced using model organisms by providing analogies to other organisms or cases, I argue that generating knowledge in animal behavior genetics involves both a horizontal process of relating the mouse to the human as well as a vertical process of generating increasingly general and risky claims about

both mice and humans that together form a *scaffolding* for supporting particular research programs. Researchers at Western are especially concerned with the configuration and the strength of the epistemic scaffolding of behavioral tests, and they employ a number of techniques to precisely configure the relationship between the mouse and the human (such as emphasizing some kinds of similarities and excluding other topics as appropriate topics for scientific investigation). I argue that these particular configurations of the epistemic scaffolding act as shared frameworks for understanding mouse modeling research, and by talking about the mouse in "non-anthropomorphic" ways researchers indicate their understanding and acceptance of these frameworks.

Chapter 3 examines the representations of the human that are embedded in the epistemic scaffolds of mouse models for behavioral disorders. Using material from the Alcohol Research Consortium and the Mouse Phenome Database, I explore how researchers use particular mouse models or methods to talk about the human that is being modeled. I argue that the human is not straightforwardly "reduced" to genes through these models, even though researchers develop these models for the express purpose of finding genetic contributors to behavioral disorders. Aspects of the epistemic scaffolds of these models have flexible metaphorical entailments that offer resources for researchers to talk about human behavior in a variety of ways: In some cases, they portray both mouse and human drinking as environmentally rather than genetically controlled; and in other cases, they use mouse models to talk about multiple motivations or factors that might affect drinking behaviors or gene expression in humans.

In chapters 4 and 5, I shift my focus from the epistemic work of animal behavior genetics to the institutional settings and political cultures in which animal behavior genetics research takes place. *Chapter 4* maps some of the different social groups, regulations, and institutional structures that shape the practice of animal behavior genetics research by describing the "career" of a single, fictional mouse as it moves through the laboratory. I

argue that the experimental mouse is not the only mouse that is being produced in the space of the laboratory, and that researchers, technicians, regulatory committee members, and the animal care staff are also producing other kinds of mice, such as the mouse that breeds, the mouse that feels pain and suffering which must be minimized and managed, and the mouse that is a commercial object. I look specifically at the care of the laboratory mouse to show that in some cases researchers talk about the institutional structures and practices that support mouse research as seamlessly aligned with experimental practice, and in other cases they talk about these practices as conflicting with or constraining mouse research.

Chapter 5 explores how animal behavior geneticists themselves talk about the political culture that they work in and how it affects their research. Public communication of their research is something that many practitioners discuss with a sense of obvious unease, and I explore how researcher's understandings of "the public" contribute to the perception of talking to the lay public about behavior genetics as an activity that is difficult and dangerous. There is little consensus amongst researchers about how best to deal with these issues, but I explore some of the strategies that researchers use to attempt to manage the public image of their field and the public understanding of behavior genetics data. I argue that the messages that researchers at Western attempt to impart to the public are similar to the messages that they impart to scientists in training and the broader behavior genetics community. When talking to the public, researchers emphasize the long timeline of behavior genetics research and the complexity of human behavior.

Finally, in the *Conclusion* I reflect on how this extended look into the world of animal behavior genetics contributes to discussions about the meaning and merits of contemporary genetic research. Examining the epistemic scaffolding of animal behavior genetics research offers a way of clarifying what is at stake in scientists' discussions about methods, and a way of conceptualizing the human that is absent but always present in the laboratory.

1 Controlling Complexity: Expectations and Experimental Practice in the Smith Laboratory

Ordinarily, the students left almost immediately after lectures finished in the introductory behavior genetics class, and I was hard pressed to catch up with them before they disappeared back into their laboratories and offices. Today, the class was unusually quiet after Laura's lecture on "environmental interactions." She presented in rapid sequence a variety of environmental factors that could change a mouse's behavior, from experiences in the womb to light levels in the laboratory testing rooms. The forcefully delivered lecture inspired a sense of awe and frustration in the aspiring scientists, who lingered in the semidarkness of the classroom after the slideshow ended. "With all of this complexity," one of the students remarked, "it's hard to feel like you have a prayer." Laura nodded, and reassured them that they were all in the same boat. Sometimes, she said, she felt depressed by it too.

By many accounts, doing animal behavior genetics research is a messy business, and not just because of the amount of mouse poop involved. Behavioral traits like impulsivity, anxiety, and shyness are ill-defined, culturally specific, and subject to historical reformulation (Frazzetto & Gross, 2007). Making tests to measure the essence of these traits is tricky, and researchers argue that tools like the Diagnostic and Statistical Manual of Mental Disorders (DSM) define psychiatric disorders in ways that are too imprecise to be useful for biological research (Gottesman & Gould, 2003). Environmental and biological

contributors to behavior are similarly challenging to nail down. Vulnerability to diseases like alcoholism is partly hereditary, but what genes a person inherits and how these genes interact with the environment to produce disease remains unclear (National Institute on Alcohol Abuse and Alcoholism, 1992, 2003). Marcus, a postdoc in the Smith laboratory, complained that it is difficult to look at how gene expression in the brain changes when animals drink because alcohol is a "dirty drug" that seems to have no specific target, affecting many different brain regions and receptors. A news article in a recent special issue of *Science* magazine on the genetics of behavior summarized some of the challenges that researchers identify in their search for specific genetic contributors to behavioral disorders:

> All we really know so far is that behavioral genes are not solo players; it takes many to orchestrate each trait. Complicating matters further, any single gene may play a role in several seemingly disparate functions. For example, the same gene may influence propensities towards depression, overeating, and impulsive behavior, making it difficult to tease out underlying mechanisms … Environment also plays a strong hand, bringing out, neutralizing or even negating a gene's influence. And genes interact with one another in unpredictable ways (Holden, 2008, p. 892).

In many cases, associations between a gene and a behavioral disorder that initially seemed promising to researchers later proved difficult to reproduce or only seemed to hold true under certain conditions or in particular human populations. One researcher quoted in the *Science* magazine article summarized the mixed evidence on a gene associated with anxiety, saying that it "leaves a trail of intriguing hints … but nothing that solidly replicates" (Holden, 2008, p. 895).

Scientists often use the term "complexity" to summarize this constellation of characteristics that they argue makes behavioral disorders difficult to study. Unlike diseases where only one or two genes are involved and the impact of the environment is minimal, researchers argue that it is likely that multiple genes and multiple environmental factors contribute to the development of "complex diseases" like behavioral and psychiatric dis-

orders. In the years leading up to and especially following the publication of the human genome sequence, much of the effort in many fields of genetic research has been redirected from studying "single gene" disorders to studying many different kinds of complex disorders such as depression, cancer, and heart disease. In the field of behavioral and psychiatric genetics, researchers argue that new methods are needed in order to better capture the complexity of the genetics of behavioral disorders (Caspi et al., 2002, 2003; Gottesman & Gould, 2003; Hamer, 2002; Kendler, 2005; Moffitt, 2005).

This chapter examines how researchers in the Smith laboratory and related laboratories at Western engage with complexity as a practical matter in the laboratory.[1] What do Smith laboratory researchers mean when they talk about behavioral disorders as "complex"? What do they see as the sources of complexity, and how do they attempt to control these sources to produce stable experimental results? How do they decide what is reasonable and practical to control, and how do they answer that question differently than other research groups? And finally, what kind of information do they hope to produce about the genetics of complex disorders? In addressing these questions, I treat scientists' discussions about the complexity of behavior as not just a description of the nature of behavioral disorders, but as a set of expectations that has particular consequences for experimental practice. In their examination of the psychiatric genetics literature, Arribas-Ayllon, Bartlett, and Featherstone (2010) argue that "complexity" can be thought of as a rhetorical formulation that does particular work for the field. They argue that descriptions of the complexity of psychiatric disorders found in review articles simultaneously provide a way of accounting for past failures in the field and of showing a new way forward. "Complexity" explains why previous research efforts might have produced inconsistent results about how genes impact particular psychiatric disorders, and it constructs careful optimism about the promises of new methodologies and what behavior geneticists can hope to accomplish.

[1] In this chapter, I talk about the Smith laboratory and other laboratories at Western that share similar views on methodological issues interchangeably.

This chapter examines how the Smith laboratory researchers talk about "complexity" as part of their day-to-day experimental work. I explore how this complexity talk is used to account for past problems, argue for particular standards of experimental practice, and manage expectations about what researchers can hope to discover about the genetics of alcoholism.

To describe how experimental work unfolds in the neuroscience department at Western, I draw on several theories in science and technology studies describing how scientific knowledge is produced. In his discussion of experimentation and replication, Harry Collins (1985) argues that so-called "crucial" experiments do not themselves decisively determine whether or not a particular hypothesis is true, because other scientists can always question whether a particular experiment was done well. He argues instead that scientific knowledge accumulates through a process of social negotiation in particular scientific communities about what should count as a well-done experiment and trustworthy data. While Collins focuses on case studies in physics that involve a heavily contested experimental result, Rheinberger (1997) points out that knowledge production does not always center around a few key experiments. Using a case study from a molecular biology laboratory, Rheinberger describes a way of conceptualizing how scientific knowledge can be produced through "experimental systems," where experimental techniques and the phenomena that they investigate are gradually stabilized and clarified over time.

I argue that both of these approaches are useful for understanding methodological discussions in the Smith laboratory and the broader animal behavior genetics community, because researchers are debating both what constitutes a well-done experiment and how quickly well-done experiments can be expected to produce scientific knowledge. Methodological discussions about how best to study complex disorders are partly about forming agreement on specific attributes a good experiment or a good practitioner should have, but they are also about forming assumptions about how stable the field's experimental systems

are and how quickly and easily researchers will be able to extract information about the genetics of behavior. In the Smith laboratory, where researchers are especially attuned to methodological issues, there is a great deal of discussion about what a "well-done" experiment looks like. In particular, experimental controls are central to the Smith researchers' conceptions of what counts as good experimental practice. Methodological dicussions in the Smith laboratory also reveal expectations about the limits of experimental control, and how much trust researchers can place in the completeness and validity of the knowledge that their experimental systems generate. I explore some of the ways that new practitioners in the neuroscience department learn about what counts as acceptable experimental practice at Western—such as "cautionary tales" that impart experimental lessons—and moments where they develop expectations about how stabilized their experimental systems are–such as "complexity crises" where students reformulate their assumptions about what kinds of findings they are likely to generate.

While much time and attention is devoted to creating controlled experiments in the Smith laboratory, not all aspects of the experimental system are seen as equally difficult to control. While researchers see the laboratory environment as something that is impossible or impractical to fully control, the mouse genome is seen as well-controlled. Social and technological committments to the stability of the mouse genome make this assumption difficult to question. Finally, I explore how the Smith laboratory researchers contrast their standards for experimental practice and expectations about the outputs of their experimental systems to those of other animal behavior geneticists. The researchers in the Smith laboratory, many of whom have backgrounds in psychology, often complain that researchers with different disciplinary backgrounds underestimate the "complexity" of behavior and regard animal behavioral experiments as either too easy or too difficult to do. These accounts of their experimental practice as different from that of other practitioners help to constitute their identity and assert their expertise as behavioral specialists within

the animal behavior genetics field.

1.1 Producing Knowledge about "Complex Disorders" in the Smith Laboratory

In *Changing Order*, Harry Collins (1985) speculates about what science might look like from the perspective of a mouse; that is, an especially smart "philosopher mouse" who is trying to figure out what rules humans use to produce scientific knowledge. Collins's rodent subject starts by defining science's ordering principle as the process of testing facts by experimentation and replication. Scientists make observations in controlled settings that can be recreated by others, and when the same phenomenon is observed repeatedly it counts as scientific knowledge. Collins points out that this description of science almost immediately encounters a serious difficulty, since no two experiments are ever *exactly* the same. Not only would it be impossible to re-create the identical time, place, and conditions under which the first experiment was conducted, but Collins argues that even if researchers could create a new experiment that was identical in every way to the original it would "amount to no more than reading the first experimental report for a second time" (p. 34). To remedy this problem, the philosopher mouse might argue that replications should be based on "approximate repetitions" that are not exactly the same, but resemble the original experiment in certain important respects (p. 29–30). Once again, Collins argues that the philosopher mouse's rule encounters a problem: Who gets to decide which respects are the important ones, and what should count as the same experiment for the purposes of replication?

Developing rules for this problem, as Collins explores, turns out to be quite tricky: Do experiments that are conducted by non-scientists count? What about experiments using radically different or even "unscientific" methods? Even if the replication is performed by another scientist using the same methods, how can one judge if the experiment is being

conducted competently? Collins argues that these questions are especially difficult to answer at the leading edge of scientific research. Established experiments produce baseline results that show that the equipment is working and that the experimenter is a competent practitioner, but when researchers are exploring new phenomena it is more difficult to tell whether an experiment is failing to detect a phenomenon because the equipment is broken or the experiment has been incompetently performed, or because the phenomenon is not there. Collins (1975, 1981) has termed this problem the "experimenter's regress." He elaborates:

> Usually, successful practice of an experimental skill is evident in a successful outcome to an experiment, but where detection of a novel phenomenon is in question, it is not clear what should count as a "successful outcome"—detection or non-detection of the phenomenon. Thus arguments concerning the existence of the phenomenon turn, not upon experimental results, but upon what comes to count as a "well-done experiment" (Collins, 1981, p. 34).

Much to the frustration of the aspiring philosopher mouse, Collins (1985) suggests that there are no hard and fast rules for determining whether an experiment has been successfully executed. Scientists may try to assess the quality of the experiments by conducting more tests to calibrate the experimental apparatus or to compare the experimental outcomes to other phenomena, but Collins argues that these tests only demonstrate the need for more tests, deepening the regress. Instead, he argues that researchers eventually overcome the problem of the experimenter's regress through a process of social negotiation that establishes what counts as a "well-done experiment." Research communities develop standards for what constitutes good experimental practice, such as who counts as a skilled technician, what counts as the same type of experiment, and what the expected outcomes of an experiment are.

For researchers at Western, one of the most important features of what makes for a "well-done" experiment is the types of controls—measures to eliminate possible alternative explanations for experimental outcomes—that the scientists have used in the experiment.

Many kinds of experimental practice incorporate controls, but the time and attention devoted to controls is one of the defining features of experimental practice in the Smith laboratory. For Western scientists, designing and executing well-controlled experiments is a hallmark of both good research and of good researchers. Chloe, a graduate student in the Martin laboratory, commented admiringly on an article that she read for the introductory behavior genetics class, exclaiming that "they controlled for things that I never even thought of!" In my interviews with researchers at Western, they frequently described themselves as "anal," "picky," or "perfectionistic" people, and expressed pride in the amount of care they took in executing experiments and the lengths they were willing to go to ensure consistency. In one case I interviewed several members of the same laboratory where each person described themselves as the most perfectionistic person in the lab. When I asked Hannah, another graduate student in the Martin laboratory, if she could think of any controls that people in her lab did that she thought were probably unnecessary, she replied:

> Probably not, because I think I'm more obsessive than the others in my lab. I can tell you one thing that they do that I don't like is when they're doing these holding cages, a lot of times as they're setting up, they'll put a home cage on top of a holding cage so they can get that next group of mice out.[2] And there's, you know, a mouse in there. Well, that's a different experience for that mouse than for the rest of the mice, because it's dark, something's on top of it, you've got something new to explore, it's more noise, it's more shaking. So I just sit there and just grimace whenever I see that.

It is not only principal investigators and graduate students who are concerned about the noise and light levels in an experimental mouse's home environment. Technicians, too, are trained to take the management of the controlled laboratory environment quite seriously. On one afternoon I overheard two technicians gossiping about a technician from another laboratory, and one remarked—to the horror of the other—that she had

[2] A "home cage" is the cage where mice live before and after experiments, and a "holding cage" is a cage that researchers use to temporarily house mice during an experiment, often after they have received some experimental treatment (like an injection) but before they are tested.

seen this technician *yell* at someone in front of the mice (her emphasis).

In some cases, the types of things that researchers at Western control for and their methods for doing so may appear quite extreme to the outside observer. One professor, for example, was convinced that the kind of plastic that the cages were made of had an important effect on mouse behavior. Dennis was skeptical of this claim, but told me that he always reports on the type of cages he uses in the methods sections of his papers, just in case. Matthew, a graduate student at Western, was worried about what he called the "weekend effect." He pointed out that the environment of the research facility is different during the week when many people are around than it is on weekends, and so he worried that if he came in to run behavioral experiments on the weekend that he might get different results. Alex, another graduate student in the department, was particularly concerned about smell. In an interview he told me that he once forgot to wear deodorant on the first day of a two week long experiment and he didn't wear deodorant for the rest of his experimental run so that the rats wouldn't sense a change in his body odor.

To illustrate some of the many different aspects of experimental design and execution that researchers at Western are concerned with, consider some of the controls that researchers would use in an experiment like the "conditioned place preference" test. In this test, which is used frequently at Western to assess a rodent's "preference" for a drug, researchers train mice to associate a distinctive space with a particular drug and then look to see if the mice return to that space (presumably because of positive or negative associations with the drug). The apparatus that researchers at Western use for conditioning mice is about the size of a shoebox and has two compartments with different floors, one wire grid floor and one metal floor with holes in it. Over the course of several weeks, researchers train mice by injecting them with drugs and then immediately placing them in one side of the box, so that the mice come to associate the feeling of the drug with the space that they are in. On the "test day," researchers put the mice back in the box and see

how much time they spend in the space that was "paired" with the drug and how much time they spend in other "unpaired" side of the box.

What kinds of controls do researchers at Western think are needed to execute a well-done conditioned place preference experiment? Many of the control measures that researchers at Western employ are aimed at maintaining consistency, so that all of the mice in an experiment have approximately the same experiences and experiments can be compared to each other. This might include ensuring that all of the mice are of the same genotype and the same sex, trained in the same conditioning apparatus with the same distinctive markings in each space, the same training schedule, and the same technique for delivering the drug. A manual from Western professor Brian McGraw's laboratory on how to run conditioned place preference experiments reminds students and technicians that "deviations should be minimized whenever possible," and instructs them to test animals in the same order, to take breaks between the same groups of mice, and to start testing at same time every day (plus or minus thirty minutes). Western researchers are also concerned with maintaining consistency in the environment of their research subjects when they are not engaged in experiments. They monitor the noise and light levels in the animal housing rooms as well as in the testing environment, and they try to ensure that mice are fed the same food, sleep in the same bedding, and maybe even are handled by the same person.

Some of the types of things that researchers at Western control for are informed by the alcohol literature (such as the established difference in preference for alcohol in different sexes of mice), and other factors are included because of their potential effects on mouse behavior (such as the light levels in the testing rooms). Still more sources of variation become visible in the process of setting up and troubleshooting experiments. Brian recalled that when he was first setting up his laboratory, he decided to put his conditioned place preference boxes inside old refrigerators to block out ambient noise from the surrounding environment. After testing this new setup, he found that he was getting consistent results

from mice on one shelf, but not on another. He eventually determined that the fan he installed for ventilation in the old fridges was blowing air more forcefully at mice on one shelf than on the other, and so he had to reposition the fan so that the airflow was consistent throughout the fridge chamber.

In addition to practices that maintain consistency within and between experiments, researchers also use "control groups" of mice in the experiment or run additional experiments to control for alternative ways of interpreting experimental outcomes. Alongside the mice who are receiving an experimental treatment, researchers run a control group of mice that receive the same treatment as the experimental mice but receive a dose of salt water in the conditioning apparatus instead of a dose of drug. When conducting experiments with knockout mice, researchers might also test a group of the genetically unmodified "wild type" mice as a control alongside the knockouts in the same experiment. Performing additional experiments to rule out the possibility of side effects that might interfere with the experiment is another way that experiments are controlled in the Smith lab. In the case of the conditioned place preference test, researchers argue that there are several alternative explanations that need to be controlled for, such as the possibility that the experimental treatment might alter the learning and memory, motivation, or locomotion of the mouse. If knockout mice divide their time equally between the two rooms in a conditioned place preference box on test day, for example, it might be because knocking out the gene makes the mice indifferent to the drug or it might be because the knockouts are unable to remember which side of the box they received drug in. Researchers might conduct additional experiments to test the mouse's ability to form memories to control for this possibility. An introductory textbook on behavioral testing recommends that researchers conduct a "mouse physical" on any genetically altered mice to check cognitive and physical functions like vision, smell, and the ability to walk, learn, and remember prior to behavioral testing (Crawley, 2007).

Finally, researchers at Western also devote considerable time and attention to managing the inevitable contingencies that arise in the course of running experiments, either by minimizing the impact of these changes on the laboratory environment or by tracking the changes where they cannot be prevented. The building-wide shift to daylight savings time, for example, was announced in the weekly Smith laboratory meetings no less than a month in advance (and every week thereafter) so that researchers could prepare to incrementally shift their mice to the new schedule. Similar effort was directed at tracking the possible effects of switching over to a new brand of bedding in one of the animal housing facilities at Western, and especially valuable mice and mice that were in the middle of an experiment were kept on the old bedding, just in case. Noisy events like construction work, clanging carts, or people talking on cell phones in the hallway were also worrisome to Western researchers. Jeffrey, one of the Smith laboratory managers, recounted that he had resorted to holding a giant pillow over the fire bell during past drills to minimize the noise when the building managers could not be talked out of testing the alarm system on the floor that housed the mice. Brian's instruction manual for conditioned place preference experiments recommends that if "personal injury" occurs during testing the researcher should should keep running the experiment but take care not to get blood on the mice.

To summarize, experimental controls are a central characteristic of what counts as a "well-done" experiment for the Smith laboratory researchers. While researchers describe many of these controls as measures that are designed to keep things the same between and within experiments, Collins's (1975, 1981, 1985) work on replication suggests that the elements of an experiment that researchers describe as "the same" are rarely ever literally identical. At Western University, the brands or models of conditioning boxes are sometimes different between or even within laboratories, the bottles of alcohol and syringes might come from different suppliers, and the experiments might be conducted by different people. Whether or not these differences are considered relevant, argues Collins,

depends on social negotiations and agreements within the research community. Within the neuroscience department at Western, there is a good deal of agreement on the types of variables that should be controlled for and the types of measures needed to control them well. In interviews, I asked researchers to name the five most important things that they control for in their experiments. When I posed this question to Dennis, he produced without hesitation a much longer list of factors that he would expect to see listed in any behavioral paper, regardless of the experiment being performed: the genotype, sex, and age of the mice; the size of the cage and the number of mice kept in each cage; the kind of ventilation system provided for the cages; food and water availability, food and bedding type; the temperature of the mouse rooms; the light/dark schedule in the rooms; and whether experiments were done during the light period or the dark. The lists that other scientists at Western produced had a high degree of overlap with Dennis's. Researchers also stressed similar ways of controlling these variables, such as handling the mice in a way that would not cause stress, or having the same person perform all of the experiments.

Researchers' standards of acceptable experimental practice are not always homogeneous, however, both in the broader animal behavior genetics community and even within the neuroscience department at Western. Alex, for example, pointed out that researchers at Western differentiate between some features of the conditioned place preference experiment that were considered "the same" in his old laboratory. In an "unbiased" conditioned place preference experiment, the grid floor and the hole floor are typically assumed to be similar enough that when a mouse is first placed in the conditioning box, it should have no preference for one floor over the other. While the researchers at his former laboratory worked from the assumption that the floors in their conditioning boxes were similar enough, he explains that Brian does not make the same assumption in his experiments:

> So what Brian does, and what took me a long time to figure out that I needed to do, but now I realize it's correct even though I don't always agree that it's correct because ideally, like I said in an unbiased design, there's no initial preference, so this should be okay, right? But the problem is sometimes some

> initial preference does creep in for whatever reason. Brian takes it one step further and compares the amount of time that the grid plus guys spent on the grid floor and the amount of time that the hole plus guys spent on the grid floor. So now it's a between groups [experimental design], you know. I didn't like it because that means I had to run twice as many animals because I'm halving my groups now, alright? If I want 12 animals per subgroup for each of my 4 groups, I need 96 animals, and I can only run 24 animals at a time, so I have to run four different experiments which take two weeks each, which basically means I'm here every fucking day.

Brian's experimental design is based on the assumption that the two floors might not be similar enough, and that in some cases animals might prefer the one floor over the other. Making this assumption means that the experiment takes twice as long to run, but Alex considers it necessary to do in order to produce "correct" experimental data.

Collins (1985) argues that these sites of agreement (or lack of agreement) in the field about the quality of experimental work are the sites where researchers ultimately break the paradox that arises from trying to use experiments themselves to test whether new experiments are working; or in other words, where researchers negotiate what counts as a good experiment and therefore what counts as scientific fact. Hannah, a graduate student at Western, explains how she might assess the value of a new paper describing an activity experiment in the animal behavior genetics literature by looking at who performed it and how it was executed. She says:

> I think you just kind of look at the quality of the behavior. You know that strain of mice and or you know those rats and you know what should be happening, and is that happening and do they have the proper control groups, do the control groups look good? If they say "sound-attenuated chamber," that gives you at least some hope that it's not in, and some people who even in our department do this, that it's not in the middle of an open lab with like people working and running gels right next to them.

Hannah starts by asking whether the research group in question has produced expected baseline results from known strains of mice or rats. Collins argues that this "test of a test" is unlikely to bring about a resolution to any serious disagreement about whether the new experimental findings are correct (although it might serve as a demonstration that the

person who performed the test is a competent experimenter). The graduate student goes on to question whether the experimenters have satisfied the requirements for what she would consider to be a well-done experiment, such as using the "proper control groups" and conducting the experiment in a sound-attenuated chamber instead of in the middle of a busy laboratory. Collins argues that settlements around aspects of experimental practice such as these in the research community are what is needed to break the cycle of the "experimenter's regress" and to establish whether this is a true experimental result.

Collins's work on experimentation and replication helps to make visible the social processes by which research communities establish accepted experimental practices, but the role that these negotiations and agreements play in the accumulation of knowledge may not be the same in all scientific fields. The case studies that he uses are centered around experiments that are designed to give reasonably definitive answers to specific questions, such as whether gravity waves exist or not. By solving the question of whether a particular research group has built a credible gravity wave detector, researchers also solve the question of whether gravity waves exist. Collins writes:

> Scientists and others tend to believe in the responsiveness of nature to manipulations directed by sets of algorithm-like instructions. This gives the impression that carrying out experiments is, literally, a formality. This belief, though it may be suspended at times of difficulty, re-crystallizes catastrophically upon the successful completion of an experiment (Collins, 1985, p. 76).

In Collins's description, the uncertainty that scientists may describe while they are setting up or troubleshooting an experiment dissipates rapidly after the experiment comes to a successful conclusion. Collins refocuses attention away from the role of the "crucial experiment" itself in establishing this certainty, but does not consider cases where scientists might not consider these experiments to be "crucial" to begin with.

Like Collins, Rheinberger (1997) also takes issue with the central role generally attributed to experiments in scientific knowledge production, but he argues that the idea of a "crucial experiment" is flawed because the outcome of individual experiments in the biological

sciences is rarely convincing enough to move research programs forward. In his study of research practice in molecular biology, he argues that "experimental systems"—comprised of practices, technical objects, and research questions—rather than experiments should be thought of as the "smallest integral working units of research" (p. 28). The function of experimental systems as Rheinberger describes them is quite different than the function of experiments. Experimental systems are not designed to produce specific information about narrowly defined questions, but to generate epistemically interesting differences that become the source material for future research. In his own words, experimental systems "are not simply experimental devices for generating answers; [they] are vehicles for materializing questions" (p. 28). In order for experimental systems to remain productive, they must be stable and reproducible enough that scientists can make some sense of the results that are generated, but not so stabilized that they can no longer produce unpredictable results. When outputs of experimental systems are so well-characterized that they are no longer surprising, Rheinberger argues that these "epistemic things" turn into "technical things" that become part of the technical repertoire for setting up new experiments (p. 29). In contrast to a model of knowledge production where theories are established or refuted based on decisive experiments, Rheinberger depicts the accumulation of knowledge through experimentation as something that takes place over a much longer period of time. He writes that it is through the ongoing process of experimental work that "the contours of what will, perhaps one day, constitute the basic concepts of a science emerge in a gradual fashion" (Rheinberger, 1997, p. 13).

Rheinberger's (1997) description of knowledge production in a molecular biology laboratory nicely captures the more tentative nature of experimental results that also exists in the animal behavior genetics laboratory. In the contemporary animal behavior genetics field, few researchers express expectations that mouse behavior will respond with "algorithm-like instructions" to their experimental manipulations, at least in the short

term. Instead, researchers argue that because behaviors are "complex," it will take longer to establish patterns of gene action in behaviors and to solidify the relationships between particular genes and particular behaviors. Rheinberger's focus on the temporality of experimental practice and the process of gradual stabilization also suggests another area where social negotations and settlements might be taking place: around the assumptions and expectations that researchers might have about the timeline of knowledge production. Researchers might disagree on whether particular experimental systems are still producing epistemic things or have already been stabilized into technical things, or how quickly they expect the contours of epistemic things to become defined and knowledge to accumulate.[3]

In the remainder of this chapter, I explore how researchers at Western University establish both local standards for *acceptable experimental practice*, and local *expectations* about what researchers should expect a well-done experiment to produce. As I have argued above, experimental controls are central to the Smith laboratory researchers' understandings of what good experimental practice looks like, and these understandings are an important part of how they evaluate the worth of experimental data produced using animal models. But it does not necessarily follow that just because the Smith laboratory researchers consider an experiment to be well-controlled they also expect it to produce definitive results. Experimental systems in animal behavior genetics may produce contradictory and confusing results that researchers interpret as being part of the groundwork for understanding the full "complexity" of behavior.

A short example will demonstrate how both of these processes are at play in knowledge production in the animal behavior genetics field. In the introductory behavior genetics class at Western, Laura opened one of her lectures with a story about a set of experiments conducted in the mid-1990s in a laboratory that she was working in. The researchers were

[3] See Mody and Lynch (2010) on "test objects" as particular components of experimental systems that can have the characteristics of either epistemic things or technical things over time or in different research contexts.

testing a knockout mouse that was missing a gene that they thought might be related to anxiety, and they ran the mouse through three different tests for anxiety: the "elevated plus maze," the "light-dark box," and the "open field test." As they were concluding their experiments and preparing the data for publication, they discovered that two other laboratories had also done experiments with knockout mice missing the same gene, using the same tests for anxiety. The outcomes of their studies, however, were quite different. While the laboratory that Laura was affliated with reported that the knockout mouse had the same level of anxiety as the unaltered mouse in all three tests, the second group found that the knockout mouse was less anxious than the wild type in the same tests, and the final group found that the knockout was less anxious in two tests, but more anxious in a third.

Did the researchers in Laura's laboratory think that the other experiments were similar enough to count as replications of their own experiments? And if so, which experiments did they judge to be well-done experiments and which ones did they think were poorly executed? Researchers' reactions to these contradictory results suggest that there are social negotiations taking place around both what counts as acceptable experimental practice and around the stability of animal behavior genetics experimental systems. Laura recalled that the editor of the journal where the three groups had planned to publish their work contacted them and asked them to consult with each other first so that they could "come to agreement" about what kinds of conclusions could be made about the role of this gene in anxiety before publication. The editor's attempt to facilitate a process of social negotiation amongst the three research groups, however, was met with surprise and even some indignantion by the authors in her laboratory. Laura recalled that she felt like they were being asked to "massage" their results. Her interpretation of the data from the three labs was not that one experiment was likely correct and the others were wrong, but that all of the results might potentially have some truth to them. She recalled that when she looked

at the papers from the other two laboratories, her first impression was that differences in the testing protocols alone were enough to potentially explain why the other groups had different findings, such as the light levels in the rooms where the research groups conducted their testing. The differences between the three research groups' studies also formed the basis for future research projects. Following the publication of their articles, two of the research groups decided to investigate the reasons behind their divergent results by visiting each other's laboratories and eventually concluded that the differences in their results were probably due to differences in the way that the mice were produced and bred in each laboratory. In this example, the social negotiation taking place around these three contradictory studies was not around whether the gene in question contributed to anxiety, but around how to stabilize "anxiety" for genetic study. While the journal editor was eager to negotiate agreement on specific facts about the genetics of anxiety, the agreement that the study authors seemed to settle on is that researchers *do not yet know* all of the factors relevant to studying the genetics of anxiety and should proceed accordingly.

1.2 Establishing Experimental Practices and Expectations at Western University

This section looks in greater detail at how standards for experimental practice and expectations about the outputs of experimental systems are established at the department of neuroscience at Western University. There are many ways in which local understandings about what constitutes good experimental practice are formed and circulated both at Western and in the broader research community, from informal conversations with colleagues to the process of reviewing papers. I will focus in particular on two sites where new members of the field are socialized at Western: the re-telling of "cautionary tales" that describe past experimental problems in the department and "complexity crises" that graduate students report during the training process. Continuing on from the last section,

I explore how the circulation of cautionary tales and complexity crises provide opportunities for imparting information on both how researchers should do experiments and what researchers should expect experiments to produce. New researchers at Western learn about the limits of control along with the importance of controls in doing experiments on "complex disorders." Through their experiences as experimenters and stories of other researchers' experiences, they develop a set of expectations about how much information about genes they are likely to find, how much they should trust their own results, how stable the knowledge they produce will be, and how quickly knowledge production in the field will progress.

The circulation of "cautionary tales" in the department (and to some extent, the larger behavior genetics community) is one of the ways that local standards for experimental practice are formed at Western. These stories use past difficulties in the laboratory as a way of highlighting particular experimental problems. One such story that I heard repeatedly during my time at Western University—in the introductory behavior genetics class, from current and former members of the Smith lab, and from Dennis himself—was the story of the disappearing drinking difference. This cautionary tale recalled a series of experiments with a knockout mouse missing a gene for a brain receptor that were conducted in the Smith lab more than a decade earlier. The mouse had been created by another laboratory, and Dennis had obtained the knockout so that his lab could test its drinking behavior. Dennis recounts:

> We were testing a knockout for drinking alcohol, and they drank a lot of alcohol compared to wild types. And we did it three times, and we got the same answer, three times. And we said ,"wow, that's pretty good." So we published it, and in the meantime, we'd gotten more mice to replace ours because they'd gotten too old to breed anymore. When we redid the drinking study, and there was no difference! So we tested them again, and we got no difference again! And we were scratching our heads trying to figure out what went wrong, and we screwed around with the parameters for a year trying to figure out what was different, and we'd get it sometimes, we wouldn't get it other times … one study we'd get it in males, the next one we'd get it in

> females … it was just really weird. And then those mice got too old and he sent us a third batch of mice, and lo and behold, boom! There it was again, just really big.

Although the first results of Dennis's study with the knockout looked quite solid, he later encountered problems replicating his initial research finding, even in his own laboratory. The same knockout mouse was later studied by researchers in other laboratories, many of whom could not find the difference in drinking that Dennis initially reported. Dennis recalls that situation was initially baffling and there seemed to be no reasonable explanation for why the drinking difference disappeared until the effect reappeared with the third batch of mice. At that point, Dennis suspected that there might be something about the way the other laboratory was breeding the mice that might account for the difference. A phone call to the other lab revealed that they were using different sub-strains of mice for creating their knockouts. While all of the mice they had received were missing the same brain receptor gene, some other genes in the mice's genome were changing depending on which mice the other laboratory was using for breeding at the time. Dennis's lab went on to demonstrate through what he describes as a "long and tedious breeding scheme" that if the knockouts were made using one sub-strain then they would drink more, but if they were made using another sub-strain the knockouts would drink the same amount of alcohol as the wild type mice.

The cautionary story about the disappearing drinking effect identifies a problem that is now known in the field as a "genetic background effect," where the end result of knocking out one gene depends on the other genes that are present in the rest of the mouse's genome. The background effect is one instance of what researchers sometimes refer to as "genetic complexity," where there are multiple genes affecting a particular behavior and one gene's impact depends on the other genes that are present. Controlling the genetic background of knockout mice is one way of reducing this genetic complexity so that specific genetic contributions to a behavior can be identified. For example, Robert Gerlai, who described

the problem of the background effect in a heavily cited paper published in 1996, wrote:

> It must be remembered that behavioral and neurobiological traits are fairly complex, often variable, and are most probably influenced by a large number of genes as well as environmental factors. In order to dissect such traits and to understand the complex web of interactions among the underlying biological mechanisms, it is crucial to provide adequate controls for as many variables as possible (Gerlai, 1996, p. 180–181).

The story of the disappearing drinking effect is one way by which researchers at Western illustrate the effect that other genes can have on the gene being studied, and why they think researchers should control the genetic background of their mice. Charles, a former student of the Smith lab who also recounted this story to me, said that these events made him and everyone else in the laboratory "acutely aware" of the importance of controlling for the genetic background, even before examples of background effects were published and controls for the effect became standard practice in the field.

The case of genetic background effects shows, as Collins might describe it, a case where the animal behavior genetics community dealt with contradictory results by negotiating what should count as a "well-done" knockout experiment. When initial experiments in the 1990s with knockout mice were conducted, there was no consensus on whether the background strain of a knockout mouse mattered for the purposes of comparing experiments. Today, the genetic background effect is a well-established concept that most animal behavior geneticists agree can and should be controlled by using methods such as inbreeding to create a uniform genetic background before testing.[4]

This story of the disappearing drinking effect has a specific message about what counts as an acceptable knockout experiment, both for researchers at Western and in the broader animal behavior genetics community. But other cautionary tales that circulate at Western have messages that are much less clear. One such cautionary tale was a story about a more recent problem encountered in the Martin laboratory that involved the disappearance of

[4] See Crawley, 2007, ch. 2 for a discussion of background effects and a bibliography of strategies for dealing with the background effect when creating transgenic and knockout mice for behavioral testing.

the "stimulation effect." Ava, a student who had just joined the Martin lab as these events were unfolding, recounts the story:

> We had this situation last year, two years ago where we couldn't get a stimulation effect when our mice were given methamphetamine. Many strains of mice will show heightened locomotor behavior when they're given a dose of meth compared to the saline animals. Well, all of a sudden, like we weren't seeing that in our animals. We'd seen it multiple, multiple times. They didn't show this acute response, and it was just kind of like what the hell is going on? I mean, I had people watch me make the drug, I had people watch me do everything just make sure that that wasn't the problem.

The result that disappeared here was not even a new experimental result, but a well-known baseline result that was part of the initial process for setting up more complicated experiments. The problem persisted for several months, and the Martin laboratory investigated many different possibilities that might explain the strange results. They had tried having different technicians run the experiment, used different bottles of methamphetamine, and (on the suggestion of the drug control authority at Western) even tested to see if the bottles of drug actually contained methamphetamine in case someone had been stealing it from the laboratory. The stimulation effect eventually returned, but Laura reports that they never did definitively establish the source of the problem.

Unlike the story of the disappearing drinking effect, in this case the relevant factors and appropriate means of controlling them were never found. As they retold the story, Laura and her grad student drew very different conclusions about the source of the problem. Ava suspected that the problem was that the bottle of drug that they were using was somehow ineffective, although Laura noted that the results of the laboratory analysis had shown that the methamphetamine was pure. Laura thought that the changes might be due to differences in the early environments of the mice that they purchased from a commercial supplier versus mice that they bred in their own facility. When I interviewed Laura about how her laboratory dealt with the disappearing stimulation effect, I asked why she thought that the place where the mice were bred might be the problem, especially considering that

other studies (such as the multi-sited study I described in the introduction, above) found no behavioral differences between mice that were raised on site versus mice that had been shipped from a different facility. She explained:

> *Laura:* That's right, for the traits that they were looking at. But you can't be sure for every trait, right? So that's another thing that we always look at, too. In a given experiment, we try very hard to use only animals that are born here, or only animals that are shipped so that we're controlling that, just in case for that particular trait it matters, because you know for all of the traits that they looked at it didn't matter, but who knows? It might matter for something else.
>
> *NN:* It could be different for a different trait.
>
> *Laura:* Yeah.
>
> *NN:* Or could it be that that's something that everyone thinks is really important, because it makes sense, but it actually turns out that it's not?
>
> *Laura:* It may not be, but you might as well control it, just in case.

Even though there is evidence that the place where mice are raised does not have an effect on the outcome of some behavioral experiments, Laura argues that this may not be the case for the particular experiments that she is running. Whether or not it actually does make a difference for her particular test is something that she may never know. With so many different environmental variables and permutations of these variables that might affect animal behavior, it would be incredibly time-consuming to examine each of these factors on a case-by-case basis.

The story of the disappearing stimulation effect shows that cautionary tales may impart lessons for researchers in training about attitudes towards control and the limits of control as well as lessons about specific aspects of how to conduct a well-done experiment. Even though Laura identifies housing environments as the potential explanatory factor in her re-telling of the story, her conclusions about the lessons to be learned seem to be less about the control of housing and more about the stance that researchers should take with regards to experimental control; namely, that more control is better, that researchers may never know if their experiments are controlled enough, and thus it is better to act as though

potential variables might make a difference and control for them anyway. Chloe, then a first-year student rotating through the Martin lab at the time that this problem was taking place, reached a similar conclusion when she told the story to me. She did several experiments to test two different stock bottles of methamphetamine to see if that might be the source of the problem, but found no conclusive results. She recalls:

> *Chloe:* It was just really weird. And I remember feeling so bad. This was my first experience in a lab here, and I was so nervous and I didn't want to mess up at all, and yeah, I just came in and was like I don't know. I'm sorry, I don't have the answer. I felt really bad that doing the study really didn't tell us anything. But what can you do? … I moved on to a different experiment in the lab, and that was that. They were going to go get [the bottle of drug] tested, which apparently didn't prove fruitful. That stinks. That's really weird.
>
> *NN:* Yeah, I think they got it tested and it was all pure drug and whatever.
>
> *Chloe:* That's really weird. It just shows you sometimes the science isn't that … sometimes things don't work out as planned.
>
> *NN:* Sometimes it's not as controlled as you would like it to be after all, eh?
>
> *Chloe:* Exactly.

Like Laura, Chloe also concludes at the end of retelling this cautionary tale that the moral of the story is that experimental systems are unpredicatable and there are limits to how much control researchers can reasonably exert.

In summary, the "cautionary tales" that circulate in the department of neuroscience at Western can have specific messages about particular experimental problems such as the background effect, but they also circulate expectations about what types of results researchers can expect from their experimental systems. The story of the disappearing drinking effect suggests that researchers need to exert a high degree of control—controlling not one but all of the genes in the genome—in order to produce a stable experimental result; and the story of the disappearing stimulation effect suggests that even a "well-done" experiment can still produce surprises.

Another way that expectations and standards for experimental practice are developed at Western is through the training that researchers receive as they are setting up experiments

for the first time. Training is an especially important site for graduate students in the neuroscience program to learn about what constitutes good experimental practice for researchers at Western. Brian explicitly uses the process of setting up experiments to teach students in his laboratory about the kinds of factors they need to control for in their experiments and which ones he considers to be more important than others. Prior to running conditioned place preference experiments, he asks his students to fill in a sheet with a color-coded grid that shows how they will divide their mice into groups for the experiment. The matrix allows students to visualize the factors that they need to control for, such as ensuring that some mice get their drugs on the hole floors and others get drugs on the grid floors to control for any potential differences between the way that mice respond to floors. Brian showed me one such matrix for a student's upcoming experiment. He describes:

> So each time they do it, they have to think through the counterbalancing. Sometimes depending on what the experiment is, it's really easy. In this particular case, this is a more complicated experiment in terms of the procedure and design. She had to counterbalance for a lot of things, and you can't counterbalance for everything. And there probably wasn't an experiment kind of on the shelf that, an experimental design that would do exactly what she wanted to do. But even so I don't … don't like to simply have them, oh yeah, you're running a two by two experimental design, here use this … Because you can't, depending on what you're doing, you know, what are the more important rules, what are the less important rules. When you can't meet all of the counterbalancing rules, why is this one the more important one than that one? So they have to understand that.

Filling in the matrix for conditioned place preference experiments teaches students to think about what things need to be accounted for in the experiment and to think about the limits of how many things they can control at once. In some cases, students find as they are filling out the matrix that it is impossible to control for everything. This predicament encourages students to explicitly acknowledge the trade-offs that they are making in controlling for some factors and not others.

Training is important not just for graduate students, but also for more senior researchers

who are adopting new techniques or tests in their laboratories. Information about how to execute particular protocols can sometimes be found in articles specifically devoted to communicating methodology (such as those in the journal *Nature Methods*), but some of it is also shared through informal conversations and interactions between researchers. For example, when the Martin laboratory experienced the mysterious disappearance of another baseline drug effect with sensitization to methamphetamine,[5] they performed a series of experiments to see whether using holding cages during the experiment made a difference in the experimental results. They found that they were most likely to get a sensitization effect if the mice spent ten minutes in a holding cage after they got the drug but before they went in the activity monitor. Although the information from these experiments remains unpublished, the Martin lab shares it with other laboratories who call to ask for assistance in setting up their own sensitization experiments. Rheinberger (1997) argues that these processes of trafficking and trading knowledge about experimental practice can create local customs that "can sweep through a whole community, even if they are by no means the only ones that work" (p. 86). This "traffic" in experimental knowledge shared informally between laboratories provides an avenue for local standards of experimental practice established at Western to travel through the wider animal behavior genetics community.

In my interviews with graduate students at Western, many recalled that the training they received in the neuroscience program at Western made them reexamine their own standards for experimental practice, and they typically concluded that the standards of experimental practice they had started with were not rigorous enough or did not take enough variables into account. Liam, a first year graduate student, says that his rotation through Brian's lab was a formative experience that shaped the way he thought about

[5] Sensitization can be thought of as the opposite of developing tolerance for a drug—instead of mice showing less reaction to the drug the more often they receive it, they show more reaction the more often they receive it. For example, mice who are given stimulants such as methamphetamine for several days in a row will be more active each time they get a dose of meth.

experimental controls. He recalls:

> *Liam:* I guess in trying to set up all of those experiments, I started to learn a lot about well, geez, all of these things really factor into the animal's behavior and you really need to control for all of it. So I guess that's where I really started to learn about that.
>
> *NN:* All these things, like you mean things from the lab environment?
>
> *Liam:* Things from the lab environment, all the different things I needed to control for, that's kind of a vague word, sorry. I don't know. Everything that's in the animal's environment, the injections, handling the mice, you know, I didn't really think about the handler effects, and so as an undergrad, I did research with a bunch of different people, and so we would all sort of take turns running some of the experiments so that we all weren't in there all the time. But I didn't really realize that maybe that's actually affecting our results because the animals are getting exposed to different things because of the handler effects. And so when I came here, I really learned how important all of these controls were, and how important it was to control all that.

When I interviewed graduate students about their research experience prior to coming to Western University, many students (like Liam, above) told me about projects that they had completed as undergraduates and thought were successful, but that they now realized were in some way "bad projects." Hannah, a student in the Martin laboratory, described an experiment that she had designed for her senior thesis where she looked at whether a drug that blocked a brain receptor could alter a rat's motivation to drink alcohol. She tested rats on a simple maze where one side had a bottle of alcohol, and the other side was empty. After training rats on this maze, she gave her rats the drug, and measured whether they went towards the arm with the bottle of alcohol or the empty arm of the maze. She told me that she realized now that this experimental design was "wrong," because she didn't control for other possible effects of administering this drug on her rats' behavior. At the time, she assumed that her experiment worked and the drugged rats were slower in seeking out the alcohol bottle because they were not as motivated to drink. In looking back at her old experiment now, Hannah says that she "just didn't know enough of the caveats," such as the possibility that the drug she gave her rats might have impaired their

movement. She now suspects that what she found in her project was evidence of the locomotor effects of the drug, not a difference in drinking.

Graduate students also noted that they became attuned to potential changes in the environment and how they might affect mouse behavior, and feel an obligation to act with these variables in mind and to try and control them even if they don't know whether changing a particular aspect of the environment will actually have an effect on their experimental results. Liam told me that he was upset to find that the light timer was broken in one of his mouse housing rooms during one of his first experiments, and the lights came on several hours earlier than they should have in the morning. Even though this change didn't seem to affect the outcome of the experiment, he said he felt that as an honest scientist he could no longer use the data because he knew that there had been a change in the housing environment. Ava told me that she noticed that her results in her methamphetamine sensitization experiments changed after she got a dog. She recalls:

> We were doing this selection for the meth sensitization, and it's a multi-day study, and you know, I still had classes, and so there were a lot of times when it just wasn't feasible for me to do it. And so [a technician] would do it, and she always got way more robust sensitization than I did. But the question was always why. And then, I got dogs! And my response came up.

The technician who was helping her run the experiments also had a dog, and she thought it was likely that the smell of dog was affecting the behavior of the mice. She admits that this explanation for the change her in data isn't exactly a "scientific explanation," but to her it seems perfectly reasonable to assume that the smell of a dog might affect the outcome of her experiments based on what she has learned about mouse behavior at Western. After her rotations through Dennis's and Laura's labs in particular, she says she now records "way more information than [she] ever used to," because she now sees many more things that can affect mouse behavior.

The training process at Western teaches students about what constitutes good experimental work at Western, but it also reformulates their understanding of the nature of

behavior and their expectations about what kind of information they can produce in their experimental work. Many students recalled that they went through what could be described as a "complexity crisis" at some point in their graduate training. They described moments where, like the graduate students in the introductory behavior genetics class I described in the opening of this chapter, they felt overwhelmed by all of the new factors that they had to account for in their experimental practice. After being inundated with information about the complexity of behavior and all of the factors that could alter behavioral experiments, students said that they felt less confidence in their ability to extract stable information about the genetics of behavior. Even though Hannah is now a senior student in the behavioral neuroscience department at Western, she says that she still has moments where she feels this way about her research:

> I still feel like I need to be like, well, we really don't know, because it's all … which I think is really just a common thing with students, seeing this complex thing where it's like, oh my God, there's nine hundred things that affect this! And those things affect each other, which then affects that, which affects this. So the whole thing is just all over the map, and you're just … confused, wondering how in the world you're supposed to say anything.

Hannah says that she finds it difficult to make definitive statements about her research, especially when talking to friends and family. Emily, a student in the Tremblay laboratory, describes a similar moment of coming to terms with the complexity of behavior. She says:

> Ruth and I joke about this all the time, because it's like every question answered presents a million more questions, and we're like, well, it's job security. And what they say to people is like, when you start doing this kind of work with complex traits … it's like you realize at some point what exactly you're getting into, and you're either going to run the opposite direction because you're never going to find an answer, or you just think … I don't know, you have to make peace with it, and be like I'm going to do what I can do to help. Because I'm not, the things that I do, I mean even if I were doing amazing, wonderful work in alcoholism it would be like 50 years until we saw anything.

Even though she thought the approach of studying human behavioral disorders with animal models had great promise, she said that she worried that the experimental tools

available to her were inadequate to really study the complex effects of genes. Emily describes this realization as a turning point in her career as an animal behavior genetics researcher, where she had to reformulate her expectations about what her contributions to the field might be. Her expectation now is that it might take her entire career to generate knowledge that would lead to something like a specific treatment or genetic susceptibility test. Emily was not alone in her expectation that it could take a long time before animal behavior genetics research could produce some definitive statements about the genetic contributions to alcoholism. Even senior members of the research field who had been working on alcohol genetics since the 1960s told me that behavioral neuroscience was still a "young field," and that treatments or genetic tests could be "a long way off."

To summarize, cautionary tales and training processes are two of the ways that understandings of animal behavior genetics experimentation are circulated at Western. These understandings may be about what kinds of things make for a well-executed experiment, as Collins (1985) suggests, but I argued that these experiences of experimental practice can also circulate expectations about the kind of knowledge that experiments produce. In other words, cautionary tales and training experiences impart information not just about the type of variables researchers should control for, but also about how much certainty researchers should place on their ability to know and control the variables that matter. Cautionary tales derived from previous experiments provide specific lessons about aspects of complexity to attend to in some cases, and in other cases offer more general cautions about the limits of control in the laboratory. Likewise, training students to see a long list of things to control for in their experimental work communicates expectations about the importance of creating controlled experiments for researchers at Western, but it can also reformulate students' expectations about the type of information that they can expect to produce in their careers as animal behavior genetics researchers.

1.3 Controlling Genomes and Environments in the Smith Laboratory

So far I have described Western researchers' attitudes towards control generally as fairly conservative, attempting to exert as much control as possible over the experimental situation. While it may be reasonable to say that researchers at Western have a general tendency to advocate for more control rather than less, their search for control in the experimental situation does not extend infinitely. There are many instances in which Western researchers' desire to exert control over the experimental situation bumps up against practical constraints, and even within the department of neuroscience at Western there is also variation in individuals' opinions about how important certain controls are and what counts as the same treatment for the purposes of comparing experiments. Hannah, for example, says that she sometimes breaks the "rule" set by her advisor that mice need to acclimate to a new room for at least one week prior to testing. She explains:

> So you know, sometimes it's four days, sometimes it's six to seven, but I'm not always following the [in a mocking tone] "they're in a new room, wait for a week," because you just don't have time to let them sit there for a week. And so it's like really, five days, seven days … does it matter? Now have I brought this up to my mentor? No. Because I can hear her being like, "that's the reason that your experiment doesn't work!"

In practice, Hannah has a greater tolerance for what counts as an appropriate amount of acclimation time than she thinks her advisor would allow. She explains her decision by pointing out that in the literature five days of acclimation time seems to be about average for the kinds of experiments that she is conducting, with some people testing as early as three days or as late as fourteen.

This section looks at some differences in what researchers at Western think is practical and reasonable to attend to in their experimental work. While there are many differences in the ways that individuals assess what is reasonable to do in their own experimental work, I suggest that there are also patterns in the way that researchers at Western assess

how controllable certain aspects of the experimental setting are. Following Wynne (1992), I argue that these differences in the way that actors treat uncertainty in different parts of the experimental system cannot simply be reduced to the amount of knowledge that they have about particular factors. In his work on environmental risk assessment, Wynne argues that differences in the way that actors asses the degree of uncertainty in a situation depends on the technological and social commitments to a body of knowledge and the degree of trust that actors place in the completeness and validity of that knowledge (see also Pinch, 1981). Where one group of actors sees a situation in terms of well-defined risks—where both the main parameters of the situation and the likelihood of certain events are known—another group might see greater possibilities for uncertainties and as-yet-unknown variables to arise. In the Smith laboratory, differences in the material and social commitments to the stability of particular components of the experimental system can be seen most clearly by comparing how researchers treat control of the mouse genome versus control of the laboratory environment. While researchers talk about the genome as generally well-controlled, the laboratory environment is seen as only partially controlled and researchers often point to environmental factors to account for unexpected differences in experimental results.

In the Smith laboratory, researchers assume that they know the important ways in which genomes vary and can adequately control for them. Inbred strains of mice, for example, are one of the most important tools for the animal behavior geneticist because they are assumed to be genetically identical. Researchers at Western treat potential sources of variation in the mouse genome not only as well-known, but also as practical, reasonable, and even necessary for researchers and mouse breeders to control. Dennis explains, for example, that when researchers are making genetic alterations to their mice (such as knocking out a gene), there are a number of tools at their disposal for ensuring that they know exactly what is in their mouse's genome:

> Well, there's … you can do it at the initial step when you make the knockout … but people don't. They take a shortcut at that step, because it's easy and it saves them about $ 500 worth of genotyping, and then what that does is introduce a second strain into the background … and then it's kind of a 50/50 mix. And then they breed back for ten generations, to get rid of that extra stuff that they put in there to save $ 500, that's the way everybody makes knockouts! But if you just did that on the background strain that you got the cells from in the first place, you'd avoid that entire problem. So it can be avoided, but almost nobody does, and don't ask me why.

Dennis's comments suggest that he thinks the necessary tools to adequately control the genetic backgrounds of genetically altered mice are already in place, and that it is reasonable and practical to use them. Researchers who do not use genotyping to test the genetic background of their mice before making knockouts are taking "shortcuts," and will have to control the genetic background at a later point through inbreeding anyway. Dennis does not comment on those who make no attempt to control for genetic background effects in their experiments, but presumably an uncontrolled genetic background would fall far short of his standards for what counts as a well-done experiment.

The idea that the genetic makeup of inbred strains of mice is both well-known and well-controlled is reinforced by strong social, technological, and institutional commitments. Inbred strains of mice are a standardized tool used by many research communities other than animal behavior genetics. Institutions like the Jackson Laboratory, supported by funds from the National Institutes of Health, helped to establish lines of mice as commonly used research tools, first for cancer research and eventually for many different biological research programs (Rader, 2004). The genetic purity of the mice produced by the Jackson Laboratory allows researchers to compare research done with the same strains of mice between other laboratories and over time. If the genomes of inbred strains of mice were not well-controlled, it could potentially affect research practice in many areas of the biological sciences and cast doubt on the stability of past research results. Events that could compromise the genetic identicality of inbred mice, such as mutations that happen spontaneously in the genome or contamination from other genomes that might

occur if mice are mislabeled or if they secretly mate together, are carefully monitored by facilities that breed mice. The Jackson Laboratory's website, for example, prominently displays information on how they manage the genetic purity of their mouse stock by periodically genotyping mice to check for mutations or contamination.The rate at which spontaneous mutations occur in the mouse genome is also seen by mouse researchers as a risk factor that can be calculated and happens at a known and predictable rate (see for example Drake, Charlesworth, Charlesworth, & Crow, 1998). Because of these strong social and technological committments to the stability of the mouse genome, challenging the genetic identicality of inbred strains of mice would be a risky claim that would affect many researchers, institutions, and established research results.

Since researchers are reasonably confident in the degree of control that they have over the mouse genome, genetics has less explanatory power for making sense of day-to-day anomalies in the lab and news about previously unknown sources of variation in the mouse genome has the potential to take researchers by surprise. When I attended the 2008 meeting of IBANGS, I was curious about one of the presenters who argued that some of the differences in anxiety that he saw in his mice might be a result of "copy number variants" of a particular gene. This was a source of genetic variation that I had never heard about before, where small regions of the genome are duplicated in a way that previous methods of genome analysis did not detect (Williams et al., 2009). I was more taken aback when the author also presented data showing that even some inbred strains of mice had these small duplications that resulted in different copy numbers of some genes even within mice of the same strain.[6] When I asked Laura after the conference about her impressions of this paper, it turns out that she was somewhat surprised by it, too. She commented that although she was used to adding the caveat that their mice were genetically identical "barring spontaneous mutations," she now had a new source of variation to consider. Although

[6] Copy number variants were detected first in the human genome (see Iafrate et al., 2004; Sebat et al., 2004) and more recently in mice (see for example She, Cheng, Zöllner, Church, & Eichler, 2008).

the researchers at Western were not resistant to the idea that their inbred strains might have new sources of genetic variations, the technological and social commitments to the genetic similarity of their mouse genomes made the possibility of other uncharacterized differences seem unlikely prior to the IBANGS meeting.

In contrast to variations in the mouse genome, variations in the laboratory environment have a lot of explanatory power for making sense of experimental troubles at Western. Researchers in the Smith laboratory, and to some extent the broader animal behavior genetics community, see the experimental environment, the mouse housing rooms, and the researchers who conduct the experiments as providing a potentially endless supply of factors that could change the way a mouse behaves. When the "Mutant Mouse Behavior Network" in Europe attempted to draft a list of important parameters that should be recorded when doing behavioral experiments, they came up with a preliminary list of over fifty items, ranging from the bedding materials used in the mouse cages to whether any of the lab members were smokers or perfume users (Würbel, 2002). In Dennis's opinion, the laboratory environment extends at least this far, and maybe even further. "A controlled laboratory environment," he argued during a talk, "starts within the animal facility and extends to the experimenter's fingertips." Moreover, because it is unclear which variables in the laboratory environment impact any given behavior, he argues that researchers "should pay careful attention to as many as [they] feasibly can."

While Dennis advises that animal behavior genetics researchers should control or record information on as many factors as possible, he also suggests that there are practical limitations on how much control researchers can achieve. Dennis explains that even he doesn't control for all of the factors that he could in his laboratory:

> I mean maybe humidity is important, you know? I don't even measure the humidity in my lab. I can guess what it is from [Western City's] weather records, but you know, I'm not going to go there. Partly that's a practical thing. You know, so you find that it's humidity. So what? What are you doing to do next? [laughs] People are not going to start changing the humidity in their

lab.

In contrast to Dennis's strong opinions on the importance of controlling for the genetic background effect, here he advocates for a much more flexible approach to experimental control. He suggests that humidity may make a difference in the way that his mice behave, but his comments suggest that he doesn't consider it practical or reasonable to control for humidity in his lab.

Researchers at Western also acknowledge that there might be unknown factors in the laboratory environment that impact mouse behavior that they may not have even thought of. While there is agreement about many of the environmental variables that researchers should control for at Western, researchers are also typically open to the possibility of new environmental factors that they have not yet considered. Some of the variables I described previously—like the concern over different types of plastic in the cages—can be thought of as somewhat idiosyncratic "personal variables" that fall outside of the accepted practices of Western researchers around which variables count most and are most practical to control. But rather than ignoring these factors or attempting to discredit those who advocate for them, most researchers at Western are willing to at least entertain the possibility that these variables may also have some value. When I repeated the story of the graduate student who went without deodorant for the week to other researchers at Western, to my surprise most people agreed that this was a reasonable precaution, although no one else admitted to going deodorant-free for the sake of their experiments.

In summary, not all aspects of the experimental systems may be seen as equally well-controlled or controllable. The Smith laboratory researchers are reasonably confident about their ability to control the mouse genome. Anomalies such as spontaneous mutations can occur, but these problems are treated as well-characterized factors with known frequencies and accepted strategies for control. On the other hand, the Smith lab researchers place less trust in their ability to fully control the laboratory environment. While there is

agreement at Western on what kinds of factors in the laboratory environment matter most for their experimental practice, the list of variables that count is both long and open-ended. Consequently, when an experimental problem pops up, researchers often first look to the lab and not to the genome for the source of the trouble. I argue that these assessments about the stability of different aspects of the experimental system are not reducible to the amount of information that researchers have about them; they are also based on researchers' expectations about the trustworthiness and completeness of that information and social and technological commitments to that information. These expectations may also be subject to reformulation over time. Debates about the significance of the background effect and more recent discussions about copy number variants demonstrate that the degree to which the genome is seen as well-controlled has varied historically.

1.4 Disciplinary Differences in Expectations and Experimental Practice

As I have argued so far, researchers' ideas about good experimental practice are intertwined with their perceptions of how "complex" the phenomena that they are studying are. While the field of behavioral and psychiatric genetics as a whole might be undergoing a "turn to complexity" since the mid-1990s (Arribas-Ayllon et al., 2010), this global narrative is contested within the field. Many of the researchers at Western recalled that they experienced this shift as a vindication of a view of behavior that they had held for some time. The real turn to complexity, Dennis quipped to me, took place many years prior when researchers took up the study of behavior in the first place. This final section looks at how researchers at Western describe their understandings of good experimental practice and expectations about the outputs of experimental systems in contrast to the broader animal behavior genetics community.

Researchers at Western often point to disciplinary backgrounds to explain why other

researchers in the field have different standards for experimental practice, expectations about the complexity of behavior, and assessments of how stabilized their experimental systems are. In particular, Western researchers—many of whom were trained in psychology programs—contrast their disciplinary backgrounds with other practitioners in the field who were trained in molecular biology. Scott, a behavior geneticist from the east coast who describes his training as very similar to Dennis's, explains how he sees approaches in molecular biology in comparison to his own background in psychology:

> The perspective from the point of view of molecular biology is sort of "one gene at a time," and you're not faced at the outset with the issue of control of particular phenotypes by multiple genes, multiple sets of genes. And so that aligns itself with the medical model that historically had tried to find a gene for a disorder, right? That contrasted very much with the field of behavior genetics as it developed, which was immediately acutely tuned in to this complex control system likelihood. Anything that we were likely to study, if it had a genetic component, would likely be controlled by multiple genes, perhaps interacting, perhaps not, who knows?

Scott sees the historical roots of the behavior genetics field in the study of the heritability of psychological traits, and argues that researchers trained in this tradition will be more inclined to think of the behaviors that they study as having a more complex genetic basis. In contrast to molecular biological approaches, which he suggests are based in the "one gene, one disorder" approach, he says that psychologists such as himself expect that finding a single gene that predicts a behavioral disorder is likely to be the exception and not the rule.

The assumption of small effect sizes in psychology is one of the ways that researchers suggest how disciplinary training impacts standards for experimental practice. If researchers assume that the effect of any individual gene is quite small, then it will be even more difficult to distinguish the influence of the gene against the "noise" from the background genetics or the laboratory environment. Although there are some instances of genes that have a pronounced effect on behavior, in general behavior geneticists expect that most of the genes that they find will explain around 5 percent of the variation for a particular

behavior (Flint, Valdar, Shifman, & Mott, 2005). William, a grad student in a laboratory in Canada, explains why he thinks that higher standards of control are needed when the effects of genes are likely to be small:

> If [one strain of mouse] is a very poor performer in a rotarod, no matter what lab you do it in, it's most likely to be a poor performer. It's not going to be this amazing performer in another lab. But it's really important when you're looking for minor differences in a knockout, like what does this gene do. It's just a bunch of small effects, like say handling has an effect. It might be a small effect, but say enrichment has an effect too and these things start acting synergistically together and then temperature has an effect and humidity has an effect and once you start adding these things together you're going to have major changes.

He argues that the behavioral differences between some mouse strains may be pronounced enough that the environmental conditions of the laboratory are less relevant. A mouse that does badly on a particular test will continue to do badly on that test, no matter what fragrance the experimenter is wearing or what the light level in the lab is. But William identifies the behavioral effects of knocking out a single gene as potentially much more subtle. He argues that controlling for "small" sources of variation is more important in evaluating knockouts since he expects that the contribution of a single gene is also likely to be small. If there are many small uncontrolled sources of variation in the laboratory, these might create enough noise that researchers will not be able to distinguish a difference in behavior that is due to knocking out a gene from a difference in behavior that is due to the combined effects of noisy environments and stressful handling.

Researchers at Western also argue that disciplinary training matters not just for the level of control that researchers employ in experimental settings, but also for the types of things that they control for. Emily, who did her undergraduate degree in a psychology program, argued that the controls that the Smith laboratory researchers used in behavioral experiments were easier for her to understand than the controls used in molecular biology assays. She says:

> Even though I hadn't worked with animals before, I was a psychology major and behavioral experiments just come naturally. I was trained to think as a psychologist, not as a biologist. So understanding why we would do a behavior, what kinds of conditions it should be done in, and stuff like that just comes so much more naturally to me than understanding what molecular controls I would need and all that. So it's been much more of a struggle to develop molecular assays than behavioral, and to understand what the molecular side means. Whereas the behavioral, I don't know, I'm just more versed in the behavioral stuff, ironically, even though I haven't done it.

For Emily, the idea that the light levels in a maze apparatus might affect the outcome of behavior seemed plausible, and she attributes this to her background in psychology. Other researchers with a psychology background pointed more specifically to Robert Rosenthal's (1966) well-known body of work on "experimenter effects" in psychology which demonstrated that even subtle manipulations to the experimental setup—such as unconscious verbal or physical cues from the researcher conducting the experiment—could affect the outcome of behavioral experiments with both animal and human subjects. In contrast, Emily found it difficult to identify the relevant sources of variation when learning bench techniques. On one occasion when she was having difficulties producing results in a Western blot, her advisor suggested that she use a fresh sample of antibody since the sample she was using had been frozen and thawed multiple times. Emily was surprised by this suggestion, and said she would not have anticipated that freezing and thawing a reagent might have such a strong effect.

Discussions about the genetic background effect highlight some of the disciplinary aspects of tensions in the field about how researchers should expect genes to behave and what constitutes good experimental practice. Many longtime practitioners of behavior genetics, especially those with backgrounds in psychology or animal behavior, like the researchers in the Smith laboratory, feared that the influx of new techniques from molecular biology in the 1990s like knockouts also brought with them a set of assumptions about how genes worked that would be misleading when applied to behavior genetics. Researchers who performed the first experiments with knockout mice often came from disciplines

like molecular biology and immunology, and these practitioners raised some eyebrows in the animal behavior genetics community for being quick to associate the genes that had been knocked out with specific behaviors without considering or controlling for possible alternative explanations. Nancy, a senior researcher at the Jackson Laboratory, recalls that the idea of a "background effect" was intuitive for researchers trained in certain traditions in genetics, but not for others. She argues that ironically the potential complication of changing genetic backgrounds was least apparent to the molecular biologists, who were actually making the knockout mice:

> *Nancy:* What's different about the knockouts and transgenics, especially the knockouts, is that when that technology came along, a lot of people who weren't traditional mouse geneticists got into the research of making them and studying them. And so they came without that traditional background.
>
> *NN:* And they came from?
>
> *Nancy:* Molecular biology? Yeah. In a lot of schools, I don't know how it is now because I've been out of school for a while, but in a lot of schools during the 80s and 90s there were molecular genetics programs, and that's what they thought geneticists were. And all they knew how to do was run PCR assays and sequence DNA and you know [laughter]. They just didn't train under traditional geneticists. There's nothing wrong with that, it's just that they're coming from a different background, with different perspectives and styles.

When recalling the early days of knockout experiments in the field, many researchers I talked to at Western also recalled that it was behavior geneticists trained in psychology who were the ones that pointed out the potential confound of the genetic background in knockout experiments and that molecular biologists were initially resistant to this explanation.

Gary, an animal behavior geneticist in Canada, also argues that whether or not the idea of controlling the genetic background seemed reasonable depends on disciplinary expectations of gene effect size. If researchers expect that the effect of knocking out a gene will be pronounced, then the idea that the genetic background might change the outcome of the experiment seems less plausible. He explains:

> If you knock out a gene that influences let's say limb development and you end up with a mouse that doesn't have any legs, okay? That's such a major developmental alternation that these background genes and modifying effects or compensatory changes or the so-called flanking allele problem is basically irrelevant. It's such a robust mutation that it doesn't really matter. But for us, behavioral neuroscientists and behavioral geneticists, we cannot think like that, we have to be a bit more worried about these seemingly minuscule, negligible genetic effects. For us, these effects are real and they're not negligible.

When the effect of a gene is large, explains Gary, researchers might be justified in ignoring changes happening elsewhere in the genetic background, but when the effect of a gene is small, the genetic background problem becomes more pronounced. Developmental geneticists may be justified in ignoring the effect of background genetics or the laboratory environment because knocking out genes in the developmental pathway typically has profound effects, but he suggests that those who deal with the genetics of behavior cannot afford to do so. The smaller gene effect sizes that he expects behavioral neuroscientists are likely to be dealing with warrant a more serious consideration of factors that are "seemingly minuscule" or even "negligible" to geneticists operating under other assumptions.

The history of debate around the background effect demonstrates a tension in the behavior genetics field around how "complex" behavior is and how difficult it will be to extract genetic information. Many animal behavior geneticists that I interviewed suggested that these disciplinary tensions were more pronounced in the 1990s when knockout techniques brought many new practitioners into animal behavior genetics, and that tensions have lessened as these new techniques have been more deeply incorporated into the practices of the field. But some residual tensions remain in contemporary practice. Some researchers voiced concerns that despite changes in the ways that behavior geneticists talk about behavioral disorders, assumptions about the underlying genetic complexity haven't really changed all that much. One of the researchers that I interviewed commented, for example, that some researchers may say that there might be hundreds of genes influencing a particular disorder but secretly hope that there are only a dozen, and another suggests that other

researchers talk about complexity in grant proposals only because they know that "simple knee-jerk determinism isn't going to cut it anymore." Other animal behavior geneticists argue that the field is still trying to proceed too quickly with behavioral testing before all the technical and conceptual issues are worked out. In a commentary article in *Nature Neuroscience*, two animal researchers argued that the field needs to *festina lente*, or "make haste slowly" (Crabbe & Morris, 2004), and that the trend towards using high-throughput behavioral tests in mice to produce "ever-faster analysis" of the genetic factors contributing to behavioral disorders was likely to slow down the field in the long run. "Doing the right behavioral test carefully," they wrote, "will be more effective than doing the wrong one, many times, quickly" (p. 1176).

At Western, researchers often paradoxically describe outsiders' thinking about experimental practice in animal behavioral genetics as both too simplistic and too complicated. Researchers alternately complain that those unfamiliar with the field who take up behavioral testing have the mistaken impression that behavior is "finicky" or a "shaky thing that is difficult to measure," and that they don't "take behavior seriously" or act as though "any trained idiot" could perform behavioral tests. These descriptions of what happens when scientists that Western researchers describe as outsiders to the animal behavior genetics field take up behavioral testing suggest that there is nothing intrinsic about doing behavioral research that produces expectations about complexity. Cambrosio and Keating (1988) argue that tacit or local knowledge has an ambiguous status in scientific practice. In their study of the development of hybridoma techniques for producing monoclonal antibodies, they found that researchers spoke about the ability to produce monoclonal hybrids as something that required knowledge of immunology and demonstrated a researcher's practical mastery of that knowledge. In other cases, researchers talked about hybridoma technology work as "mere technique," something that researchers should be able to do "without necessarily coming to grips with the theoretical principles that

underlie it" (Cambrosio & Keating, 1988, p. 251). While behavior geneticists at Western treat local knowledge about how to perform experiments as a source of lessons about the "complexity" of behavior, Cambrosio and Keating's study suggests that it is also possible to treat experiments that require substantial amounts of tacit knowledge to execute as poorly stabilized experimental systems, or even as fully stabilized tests that are ready to be integrated into new experimental systems.

The conflicting descriptions from Western researchers of how animal behavior genetics tests are taken up into new communities of practice suggest that there are differences in the perceived stability of animal behavior geneticists' experimental systems. Dennis, for example, argues that researchers who take up behavioral testing tend to treat the tests as too stable. He says:

> The majority of people using these behavioral tests aren't … they're biochemists or molecular biologists, or neurophysiologists or straight geneticists of some sort, so to them it's an assay. No different than a Lowry protein assay. Except … that it kind of is! If you do it this way, you get one answer, if you do it a little differently, you're going to get a different answer.

Dennis differentiates between behavioral tests and other kinds of established tests like the Lowry protein assay that is used to determine the level of protein in a solution. While some researchers might argue that behavioral tests are different in kind than other assays because behavior is more complex, others argue that most behavioral tests are simply not ready for widespread use yet. Important parameters like the effects of noise, light levels and apparatus design have not been fully explored by the behavior genetics community, and using these tests before the operating parameters have been fully explored is likely to produce inconsistent results. Using Rheinberger's (1997) terminology, we might say that some behavior genetics researchers are objecting to outside researchers who treat behavioral experiments as stabilized "technical objects" that can be unproblematically inserted into new experimental systems instead of productive "epistemic objects" that can still produce unexpected results.

On the other hand, animal behavior geneticists' comments about the uptake of their tests also suggest that some outside researchers might regard behavioral experiments as not stabilized enough, especially if they are not used to controlling for the types of environmental variables that form the core of behavior geneticists' program of control. Rebecca, an animal behavior geneticist who often collaborates with researchers from other disciplines, says that she finds that newcomers to the field are often overwhelmed by the number of things that she tells them they need to pay attention to when conducting behavioral experiments. She comments:

> I suspect that behavior seems very complicated, and people who've never done it before try to sit down and learn how to do an experiment, it sounds to them like we're telling them too many variables that they have to pay attention to, and they get frustrated and walk away. But I don't know. I mean, I've gone in to try to learn electrophysiology, and I have the same reaction. To me it's way too difficult for me to even try to get into.

Rebecca's comments suggest that researchers in training from outside disciplines may also experience "complexity crises" when learning how to do behavioral testing, but may ultimately choose to walk away from behavioral research rather than making peace with complexity. Rebecca points to her experiences trying to learn new experimental techniques from other fields to suggest that behavioral testing is no different than other kinds of experimental practice. But for those who have no experience with animal behavior, the idea that changing the light levels or having music playing in the laboratory might affect the outcome of the test may suggest that behavioral testing is too difficult, or alternatively that behavioral tests are simply not very robust.

Western researchers' demarcations of who is an outsider in the field and what the differences are in the way that these outsiders approach behavioral testing reveal some of the variation in both experimental practice and assumptions about what experimental systems can produce in the behavior genetics field. Animal behavior geneticists at Western primarily talk about the difference between themselves and others in terms of

disciplinarity, but I have characterized the differences here as primarily about expectations and experimental practice. The debate around the relevance of the genetic background effect reveals different assumptions about the nature of gene action, and whether the consequences of knocking out a gene are likely to be pronounced enough that researchers can safely ignore other genetic and environmental factors. Discussions around the uptake of behavioral tests into new fields also reveals different expectations about the kind of data that experimental systems should produce. While some researchers appear to treat behavioral tests as stabilized technical objects that can be used to produce solid genetic knowledge, others see the data generated by behavioral testing as far too variable.

1.5 Conclusion

Near the end of one of my interviews with Emily, a grad student in Ruth's lab, the conversation had taken a rather depressing turn. Her lab had identified a promising region of the mouse genome that seemed to influence alcohol withdrawal, but their attempts to pinpoint a specific gene had not been going so well. There were too many genes in the region to say with certainty which one might be creating the effect, and when they inserted one of the genes into another mouse strain the results were inconclusive. She explained to me that these were the kinds of problems that behavior geneticists were likely come up against because they study complex traits. I asked her why she wanted to become a behavior geneticist, if this was the case. Emily replied:

> You know, behavior genetics I think is … it's really such a wonderful and interesting way to study human behavior. I think we are still, it's still in its infancy, because it's limited by both our knowledge of genetics and our ability to really understand the genome. But I think it's absolutely the right way to start looking at these issues and to tackle the complexity. You can get easily overwhelmed by the complexity, but the bottom line is the complexity is the reality. And if we really want to help people with schizophrenia and with alcoholism, we need to start tackling the whole picture.

Emily's comment nicely summarizes the interplay between expectations and experimental practice that I have described in this chapter. I have argued that when the researchers at Western talk about "complex disorders," they are referring to a particular constellation of understandings about the way that genes work to produce behavior: That most behaviors will have many genes that can influence their presentation, that the effect of these genes is likely to be small and can be influenced by other genes, and that environmental factors are also important in determining behavior. This understanding of behavior as a complex entity is strongly tied to a specific style of experimental practice at Western that is characterized by tight control over the mouse genome and the laboratory environment. For the aspiring researcher like Emily, the daily repair of breakdowns in behavioral testing and past stories about experimental practice lend a specific and local sense of the "complexity" of behavior. The experience of seeing a mouse's behavior change unexpectedly, presumably because of changes in the laboratory environment, reinforces the idea of the importance of creating controlled laboratory surroundings, but also reinforces the idea that even behaviors that are under some degree of genetic control can be modified substantially by changes in a mouse's early history and environment.

Collins (1985) argues that social negotiation within research communities about what constitutes good research practice is important for the accumulation of knowledge in scientific fields because it allows researchers to make results comparable with each other and decide what should count as scientific fact. I have argued here that negotiations and agreements about what constitutes acceptable experimental practice are also important for understanding the day-to-day work in the animal behavior genetics laboratories at Western, but that these agreements about what counts as well-done experimental work do not necessarily settle questions about whether the information produced by those experiments should be counted as stable and trustworthy. Many animal behavior genetics researchers who consider the field to be "still in its infancy" regard as premature the comparison of

experimental results to see if they replicate or refute each other. As Emily's comment demonstrates, standards for good experimental practice are important for researchers at Western not because they provide a basis for future replications, but because they establish a foundation of well-characterized experimental systems that will allow information about the genetics of behavior to emerge gradually through future experimental work. Rheinberger's (1997) work on experimental practice in the biological sciences suggests that knowledge can accumulate not only through a cycle of experimentation and replication but also through a more gradual process of stabilizing technical objects and the information that they produce. Drawing on Rheinberger and Collins together, I argue that communities of research practice can engage in social negotiations not only about what counts as a well-done experiment, but also about how much and how quickly a well-done experiment can produce.

Expectations and understandings of acceptable experimental practice are not always uniform within the animal behavior genetics field, or even within the laboratories at Western. While the Smith laboratory researchers attempt to control both the genetics and the environment of their mice, they hold different expectations about how controllable each of these aspects are. Researchers are confident in their ability to control for the important sources of variation in the mouse genome, but see the laboratory environment as full of variables that may or may not make a difference. In the view of Western scientists, other researchers in the field make even more diverse assessments of how well-controlled experimental systems are and how well-controlled they need to be. In the broader animal behavior genetics community, Western researchers' assertions of their understandings of proper experimental practice and appropriate expectations about gene action help to establish their own expertise in the field.

2 "Don't Anthropomorphize!" Linking the Mouse and the Human in Animal Models

"Step one: *Don't anthropomorphize!*," instructs the opening paragraph of a textbook chapter on animal models of psychiatric diseases (Crawley, 2007, p. 227). "Emotions are personal, internal, and highly species specific," continues the author, "[t]here is no way for a human investigator to know whether a mouse is feeling afraid, anxious, or depressed." As I read Crawley's introductory textbook on how to do behavioral research with mice in preparation for my visit to the Smith laboratory, this cautionary phrase stuck in my head. It seemed to be a reasonable caution—it is difficult enough to capture the subjective experience of another person, let alone a mouse—but also a somewhat strange one to offer as an introduction to the world of animal modeling. Using animals to study human psychiatric disorders appeared, at least to me, to be a rather fundamentally anthropomorphizing enterprise: Research in animal behavior genetics imbues mice with the power to stand in as "models" for diseases which even some researchers themselves consider to be uniquely human. At the very least, research with mice in behavior genetics is highly anthropocentric, since it is the connection between the behavioral tests administered to mice and behavioral disorders in humans that animates research in the field. Researchers who use animal models are not interested in "solving the rat cocaine problem," as one Western grad later joked to me, or perhaps even in whether mice are actually feeling anxious. Rather, they are interested in generating information that will lead to a better understanding of or treatment for humans who suffer from these disorders. Since the practice of animal

behavior genetics revolves around the connection to the human, I wondered what exactly researchers meant when they cautioned newcomers to the field not to anthropomorphize their furry research subjects.

Making sense of this apparent contradiction was even more difficult after arriving in the Smith laboratory and observing the many instances in which researchers speculate about what the mouse is doing, explain things from the perspective of a mouse, or even act out the mouse themselves in their day-to-day work. When describing the movements that mice make in mazes, researchers act out these mouse behaviors using their own bodies, peeking their heads around imaginary corners and elongating their necks and backs to create a resemblance to what they call a "stretched attend posture." More than once researchers taught me how a particular test worked by inviting me to take on the role of the mouse in the experiment. In a session where I watched a graduate student train mice to press a lever that would give them sugar water, he explained the procedure by telling me how the study would unfold from the perspective of a mouse in the training apparatus, and even held out his clipboard as an imaginary lever for me to tap as he went through the experimental design. In one of the first interactions that I had with Dennis, he described one of the tests for measuring intoxication in mice to me as follows:

> In humans, they have their own little set of assays, they make you close your eyes and try to touch the tip of your nose, and that turns out to be hard to do when you're intoxicated. They have you try to walk a line without stumbling, or talk without slurring. We've never seen slurred squeaks in the mice [laughter], but we do make them walk a line. We put them on an apparatus that's just like the balance beam you see in the Olympics.

Dennis's description creates anthropomorphic parallels between the mouse world and the human world, drawing on everyday objects like Olympic balance beams and commonsense understandings about what people do when drunk. The links between the human and animal become obvious because Dennis has aligned the mouse tests he performs in a way that overlaps with my own presumed cultural knowledge of how intoxication is measured

in humans. Researchers routinely use these kinds of constructions that blend together human experiences of alcohol drinking with the actions of mice in the lab, talking about mice that "don't drink enough to blow over" or mice that are "too drunk to drive home." Researchers also use these kinds of anthropomorphic modes to discuss experimental problems amongst themselves. When setting up a new experiment, it would not be uncommon to find some of the Smith laboratory members huddled around a prototype apparatus on a laboratory bench, with some commenting, "I think this needs to be raised, because they'll probably try to chew on that," and others taking on the role of the mouse, offering, "Well, I'd try to escape through here …"

In the previous chapter, I looked at the expectations that the Smith laboratory members have about genes and behavior, such as how "complex" behavior is, what types of things can influence it, how researchers can manage these factors, and what kind of information about behavior they can reasonably expect to produce. I focused on researchers' attempts to control mice and their surroundings to produce reliable, repeatable measurements of mouse behavior, but I left aside questions about how researchers link these data back to the human once they have it. In this chapter, I explore the question of how animal models model in behavior genetics, looking at how members of the Smith laboratory position the mouse and the human in relationship to one another in their work. To describe the relationship between animal models and human disorders, I develop the concept of an *epistemic scaffold* that involves both a horizontal process of *linking* specific information from the animal and the human and a vertical process of building increasingly broader and riskier claims about their relationship. Using the example of a behavioral test called the "elevated plus maze," I describe the different ways that researchers build, combine, and break down epistemic ladders to establish the test as a valid model of human anxiety. Researchers at Western are especially concerned with regulating the length and strength of epistemic ladders, and they employ a number of techniques like using cautious language

and identifying certain claims as inappropriately anthropomorphic to precisely configure the relationship between the mouse and the human. These techniques help researchers manage specific tensions about making claims in the field of behavioral genetics, especially around who can make a claim about genes and behavior and how much can be claimed. By talking about the mouse and the test in specific, "non-anthropomorphic" ways, researchers demonstrate both their knowledge of mouse behavior and their participation in a shared framework for understanding the relationship of mouse models and human disorders.

2.1 Animals as Models

Animal models are comprised of several different components, including the animals, the equipment, the measures, and the theories that come together to make a model complete. The model organism of choice for many researchers at Western University, the mouse, is arguably the most widely used organism in contemporary biomedical research. The mouse is what Kohler (1994) calls a "standard" organism: Like other organisms such as *Drosophila* and *C elegans*, the mouse has become a standard for biological research in the sense that it is "the thing that everyone uses" (p. 14). The intensive focus on a select group of organisms by many groups of researchers has transformed these creatures into tools that can be useful for studying many different questions and for working out general biological principles. The fly was extensively studied as a model for genetic inheritance; the fly research community used the organism to elucidate general mechanisms of heredity, such as the idea that specific physical traits could be linked to specific locations on chromosomes.

The term "model organism" focuses attention on the centrality of the organism in experimental practice, but more than just the creatures themselves are needed to make an organism into a productive laboratory tool. As Kohler argues, "[a] bottle of flies is not of much use for experimental production in itself, but only as part of an assemblage of material instruments, standard recipes and procedures, and working relationships"

(1994, p. 8). To make the fly useful, researchers needed many other resources: genetic maps that showed where other traits had been linked to chromosomes, strains of flies with known mutations, and techniques for creating new mutants, to name a few. Creager (2002) points out that researchers speak not only of model *organisms*, but also of model *systems.* Like Rheinberger's (1997) concept of the experimental system, the concept of a model system acknowledges the network of resources that are associated with research organisms. Dennis, like many researchers who are interested specifically in the genetics of behavior, has worked almost exclusively with the mouse during his career because of the availability of inbred strains and genetic maps, which give researchers doing genetic research with mice a significant advantage over researchers working with other model organisms like the rat.

Along with standardized model organisms like the mouse, researchers also need theories, equipment, and experimental protocols to make models for particular human disorders. The terms "operant self-administration model," the "drinking in the dark model," or "models for anxiety" identify sets of tests and standardized practices that model particular kinds of behavior. Each of these models involve relatively standardized protocols for conducting and analyzing the experiment, such as schedules for when mice should be given bottles of alcohol at a specific time during the dark period of its day and how researchers should measure how much the mice drink. Models for particular behaviors may require specialized equipment for conducting or measuring the experiment, such as an operant box that has levers that mice can manipulate, lights and sounds to cue particular behaviors, and a chamber for delivering "reinforcers" like food or water. Models also imply particular kinds of experimental subjects. With different kinds of equipment, operant self-administration experiments could be used on many different types of model organisms or even humans; while the drinking in the dark model was designed specifically to take advantage of the nocturnal tendencies of mice, who tend to eat and drink more

in the dark hours when they are normally active. Finally, models encapsulate specific theories about how behavior works or how animal tests are related to human conditions. The operant self-administration model draws on psychological theories about operant conditioning and the "reinforcing" properties of certain substances to train mice to press a lever to get alcohol and to determine how many times mice are willing to press that lever to get alcohol.

Together, these combination of standardized methods, instruments, materials, and theories create what Joan Fujimura (1996) refers to as "theory-methods packages." Theory-methods packages draw together specific materials (like inbred mouse strains and operant boxes) with specific theoretical concepts (like operant conditioning and positive reinforcement) in ways that facilitate research on genetics and drinking. Operant self-administration models for assessing the rewarding properties of alcohol draw together instruments for performing the experiments like operant boxes, computer programs for measuring when and how often they drink, schedules for training experimental subjects to press a lever to get alcohol, and theories that allow researchers to interpret data and relate it to more general ideas about the reinforcing properties of particular substances. Fujimura argues that the packaging together of these components promotes the circulation of particular kinds of experimental practices because it allows researchers to easily pose specific types of questions. If researchers want to ask whether knocking out a particular gene will make mice more or less willing to work for alcohol, they can choose from existing models, such as operant self-administration procedures, to use in their experiments instead of building their own model system from scratch.

Readers unfamiliar with animal behavior genetics might find themselves confused by the discussion of animal models for human behaviors so far, since these theory-methods packages look less like models that explain or represent a particular behavior and more like assays that test a particular behavior. Indeed, animal behavior geneticists themselves often

use "model" interchangeably with other terms like "protocol," "paradigm," or "test." Several animal behavior geneticists commented that the use of the term "model" sometimes creates confusion when they talk about their work to those outside of the field because "modeling" implies a very different set of practices in other scientific fields than it does in animal behavior genetics. Dennis points out that this can be a problem even within the field of behavior genetics, where the type of modeling that human behavior geneticists do is quite distinct from the type of modeling that he engages in. He explains:

> "Model" to me means something different than a model does to a statistical geneticist who's trying to develop an analysis of genetics and environmental factors like social class, education, health, illness, life stressors, and that kind of stuff. Those guys do modeling with multivariate statistics. What I mean is, I want to understand what alcoholics report, that when they're in withdrawal, after they've stopped and are partially drying out, they become anxious and that anxiety is a part of their relapse process because it drives some people to self medicate with more alcohol.

The multivariate statistical models used by human behavior geneticists that Dennis describes are the type of object that might be more easily recognizable as models in the colloquial sense of the term. These models link multiple components like education, health, and genetics together in statistical relationships, and researchers evaluate how well the models fit with available data from human populations. Dennis describes modeling in animal behavior genetics as a different kind of activity. He gives the example of how he might investigate a particular relationship between anxiety and drinking that has been observed in humans by using animals to study that relationship in a controlled setting.

A rich literature exists in the philosophy of science literature on the first type of model that Dennis describes—models that provide abstract representations, descriptions, or theories of natural phenomena (Hesse, 1966; Cartwright, 1983; Morgan & Morrison, 1999). Philosophers have described a wide array of these types of models (such as physical objects or equations), and the various roles that models can play in scientific knowledge production (for a review see Frigg & Hartmann, 2006; see also Chadarevian & Hopwood,

2004 on physical models). But some have argued that while this view of modeling as a process of abstraction, representation, and theory building works well for describing knowledge production in fields like the physical sciences, it does not always fit easily with the ways that knowledge is produced in the biological sciences and in particular with how knowledge is produced using model organisms (Ankeny, 2000; Keller, 2000; Rheinberger, 1997). Although organisms such as mice are also described as "models," these organisms are not abstract representations of humans; they are organisms with their own biology and natural history that some might argue is nearly as complicated as that of the humans that they model. As Creager et al. (2007) put it, animal models "are not simpler biologically than the humans that they illuminate by analogy" (p. 7), although they do offer many advantages as tools for studying biological principles.

Several scholars offered different ways for thinking about the way that animals models model as a process of "extrapolation," providing "exemplars," or "case based reasoning." Creager (2002) argues that model organisms model in two distinct ways: By acting as standard organisms that coordinate research on particular topics in the way that Kohler (1994) describes, and by acting as "exemplars for studying and understanding other entities and organisms" (p. 5). Creager explores how researchers studying other, less-understood viruses, like the polio virus, were able to draw on experimental techniques and concepts developed through work on the tobacco mosaic virus to further their research. Likewise, tobacco mosaic virus researchers made productive analogies to work in other fields, such as techniques for making protein crystals that helped researchers isolate the tobacco mosaic virus. To describe how researchers using the model organism *C elegans* produce knowledge, Ankeny (2007) draws on the tradition of "case-based reasoning" in medicine. Work with model organisms, she argues, can be thought of as a process of making comparisons between individual cases as a means of developing more general principles. Individual worms can be compared to other individual worms until a general

concept of the "wild type" worm emerges, and mutant worms can be compared against the wild type archetype to gain insight into how they differ. Schaffner's (2001) description of how model organisms model in behavior genetics is grounded in a similar vision of making comparisons between specific cases. He describes animal modeling as a process of "extrapolation," where researchers first establish the biological mechanisms that are at work in animals and then use this information to infer what might be happening in humans, perhaps with the secondary goal of establishing more general biological principles. The process of extrapolation presumes that mice and humans are linked through evolutionarily conserved genetic sequences or mechanisms. Schaffner is somewhat skeptical, however, about how much homology exists between humans and models, especially in the case of more distantly related models like *Drosophila* or *C elegans*.

These descripions of knowledge production using animal models portray it as a process of using better understood and experimentally tractable organisms to establish information that can be used to understand features of organisms that are more difficult to study, such as humans. In these descriptions, animal models do not necessarily have to be simpler than or exactly similar to the objects that they model, but they need to share some key features for the comparison to seem plausible. Some of the shared features that biologists identify in the mouse and the human are likely quite familiar to many readers. Researchers frequently draw on genetic comparisons between mice and humans to explain why the mouse is a suitable model, pointing to evolutionarily conserved sequences, the number of homologous genes, the similar functions of many proteins produced by these genes, and even the similar sizes of the genomes to draw links between humans and the mouse. Conserved genetic and molecular mechanisms across the species provide the basic rationale for engaging in biomedically oriented research with model organisms, especially genetic research with model organisms, but how do researchers make arguments for the particular concepts, instruments, and methods that they use to study human disorders?

What makes a mouse in a maze a good measure of human anxiety in the eyes of animal behavior genetics practitioners?

To further illuminate the ways in which researchers build plausible relationships between rodents and humans in animal behavior genetics, I describe this process as one of building *epistemic scaffolds* that connect the mouse and the human.[1] This concept draws from earlier work in science studies that describes processes of scientific claims making, especially Pinch's (1985) concept of the "externalization of observation," Latour's (1987) description of how claims are "stacked" to create claims with a higher level of induction, and Pickering's (1995) concept of "representational chains." All of these authors describes a process where scientists build on a particular observation (or "specific material capture," in Pickering's terminology) to generate claims that are increasingly removed from the original observations and are more general in scope. Pinch (1985) calls this the degree of "externality" of the claims being made, which might be thought of as the distance between the observational data and the claim. In the case of solar neutrino physics, Pinch argues that the claim that a scientist recorded "splodges on a graph" has a low degree of externality, while the claim that these "splodges" represent the production of a particular ion has a higher degree of externality, and the claim that the production of this ion indicates the presence of solar neutrinos has the highest degree of externality to the data. Experimenters must decide, he argues, on the strength of the claims they can make about their data. Making claims with low externality is probably safe but also contributes relatively little to the field, while claims with high externality are likely to be more profound but also open up more avenues for criticism from colleagues. Latour (1987) describes a similar process of "fact-building" where scientists layer claims on top of each other to create claims with higher degrees of "induction." He argues that scientific papers aim for increasingly more general and risky claims, turning data from three hamsters' kidneys into claims about

[1] Thanks to Angie Boyce and Harald Kliems for their help in naming this concept.

kidney structures in mammals more generally (Latour, 1987, p. 51). Building on these concepts, Pickering (1995) emphasizes the ambiguity that scientists face in the process of constructing claims. The relationship between experiments, facts, and theories may often appear obvious in retrospect, but Pickering argues that in most cases there are many conceptual elements and many ways in which these elements could be brought together, complicating the process of theory building. Pickering draws attention to the work involved in building "representational chains" and the choices that scientists make about how to link together these material and theoretical elements to generate new insights.

These theories nicely capture the process of relating experimentally tractable organisms to the organisms of interest that is central to generating knowledge with animal models. In this chapter, I argue that building the epistemic foundations of animal models for human disorders involves both a vertical process of making more or less general and risky statements about the knowledge that can be produced using animal models, and a horizontal process of making *links* between the animal and the human. I argue that claims about the utility of animal models are developed by continually referencing the human and the disorder being modeled. From the many similar and dissimilar features that mice and humans share, researchers select specific pieces of information that can be linked together to create convincing foundations from research programs. The resulting arguments that support the validity of animal models for human disorders resemble ladders rather than chains, in that claims about the general applicability of animal models are supported by a series of more or less risky links between the mouse and the human. Different kinds of arguments make different kinds of ladders joining the mouse and the human, which can be assembled into an *epistemic scaffold* that supports research programs on particular topics (figure 1).

Like fact-building, building epistemic scaffolds can be thought of as a bi-directional process where researchers can attempt to both build up claims and knock them down.

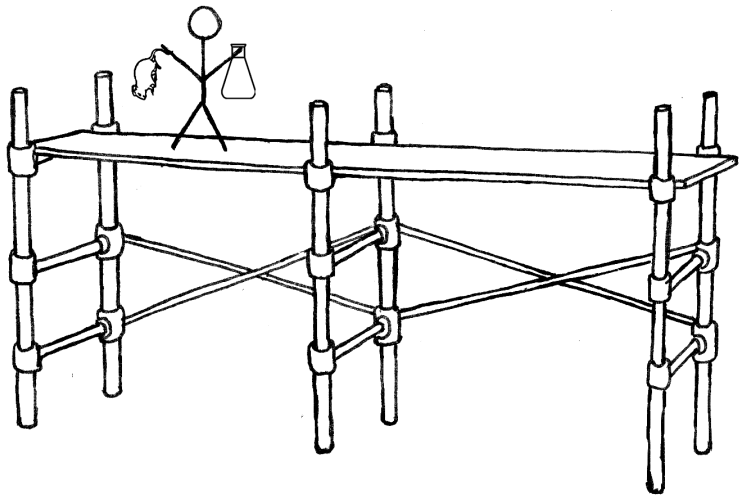

Figure 1: Illustration of an 'epistemic scaffold' supporting animal behavior genetics research. Illustration: Nicole Nelson

Both directions of the scaffolding may be subject to attempts to built stronger links or attacks by critics. Critics may attack the vertical dimension of epistemic scaffolds (for example, by suggesting that particular models don't really measure "anxiety"), or they may argue against particular horizontal links between the mouse and the human (for example, by excluding the animal mind as an appropriate topic for scientific investigation). Epistemic scaffolds as a whole can be strengthened or are susceptible to attacks in either or both of these dimensions.

This chapter pays particular attention to who is building up or taking down epistemic scaffolds, and the reasons that they offer for doing so. The processes of claims making described by Latour (1987) and Pinch (1985) have been presented as processes taking place within a sphere of professional competition, creating the impression that scientists are primarily concerned with building their claims up and breaking those of their colleagues down. Latour, for example, suggests that successful scientists aim to "prove as much

as [they] can with as little as [they] can considering the circumstances" (1987, p. 51), attempting the largest and wildest claims possible that will still pass muster with their colleagues. The downward process of "fact-breaking" takes place under pressure from other scientists or critical outsiders, who are presumably also seeking to advance their professional interests. Pinch (1985) points out that there are several reasons why scientists might want to produce claims with higher degrees of externality (such as the pressure to show a good return on the money invested in a scientific project), but he is less specific about the reasons why scientists might want to decrease the externality of their claims—he mainly points to the risk of exposing themselves to criticism from others in the scientific community.

By taking a broader view of professionalization that includes competition between professions, the credibility of professions in society (Abbott, 1988), or even non-scientific concerns, I aim to highlight some circumstances where it might be advantageous for researchers to de-emphasize the degree of relatedness between the mouse and the human or the generality of their findings in their *own* work. Rader (2004), for example, suggests that ethical concerns factor into the portrayal of the mouse's relationship to the human. Mouse models must be made to appear enough like humans so that they are a plausible substitute for us, she argues, but not so much like humans that their use in research becomes ethically problematic (p. 22). In animal behavior genetics, making assertions about the similarity between animals and humans may help to build support for their research programs, but assertions that are too broad or risky might also damage the credibility of the field in the eyes of other important actors, such as funding agencies or "the public." This chapter describes instances where practitioners attempt to increase the credibility of their profession by breaking down specific links between the animal and the human or pushing back on the level of generality of the claims being made in order to create more stable epistemic scaffolds that can resist criticisms from many different actors.

Epistemic scaffolds can also be thought of as structures that have a tendency to become "black boxed" over time. (Latour, 1987) In the development and validation of a new test, animal behavior geneticists are actively engaged in the process of creating new links that join experimental data from a new test to a human behavior, or existing information to a new theory of behavior. But in many other cases researchers are working with standardized models whose relevance to the human condition has been established through experimentation and argumentation that took place decades earlier. In the methodologically-focused Smith laboratory, however, the epistemic foundations of behavioral tests are often still visible because researchers are actively actively engaged in "test refinement" activities that make and manage links between existing bodies of experimental data and concepts of the human disease. In the next section, I will look in detail at one such test, the elevated plus maze. This test has been in use in the field for over two decades, and is widely used by researchers at Western and elsewhere in the animal behavior genetics community. In the next section, I look at how these existing connections between the mouse and the human are retold, managed and maintained in the setting of the Smith laboratory and by other animal behavior geneticists who are interested in methodology.

2.2 Building Epistemic Ladders to Support the Elevated Plus Maze

The elevated plus maze, as its name suggests, is a simple maze in the shape of a plus sign, elevated about half a meter off of the floor (figure 2). Two of the arms of the plus are enclosed by high walls (which are sometimes made of an opaque material), and the other two arms are open (which sometimes have a small lip around the edge of the arm). The arms of the maze that are surrounded by walls are called the "closed" arms of the maze, and the arms of the maze with only a small wall or no wall at all are called the "open" arms of the maze. At the beginning of a testing session, a mouse is placed in the center of the maze

where the open and closed arms meet and is allowed to explore the maze for a short period (typically between 5 and 15 minutes). Using a video camera mounted above the maze, researchers track how much time the mouse spends in the closed arms of the maze and how much time it spends in the open arms, along with a few other key measurements, such as how many times the mouse enters and exits the open arms. Depending on the laboratory, the mouse's behavior may be scored manually by a trained observer watching the video or automatically by using computerized video tracking software, or some combination of the two. These measurements, especially the amount of time that the mouse spends in the open arm of the maze, form the basis for assessing the anxiety level of the mouse.

This apparatus builds off of work with other rodent mazes in the 1960s, but it was first described as a "novel test of anxiety in the rat" in the mid-1980s (Pellow & File, 1986). It was quickly adapted for use in mice, and researchers who use the elevated plus maze generally cite a paper published by Richard Lister (1987), an intramural researcher at the National Institute for Alcohol Abuse and Alcoholism, as the first paper outlining the use of the test in mice. Lister described the technique and presented data from experiments with several different types of drugs to show that the test could be used successfully for testing anxiety in mice as well as rats. The elevated plus maze, he concluded, was an effective test that had several advantages over tests for anxiety that were available at the time. Some of the existing tests required long training periods, and many others used external stressors, such as shocks, bright lights, or loud tones to create anxiety. Not only did the elevated plus maze measure anxiety without using any of these negative stressors, it required no training period and only needed to be performed once to get results.

Today, the elevated plus maze is widely used in both academic research communities and pharmaceutical companies (Dawson, Fellow, Tricklebank, & Director, 1995), and researchers often refer to it as the "gold standard" of anxiety tests (see for example Crawley, 2007, p. 262). It would be rare to see a paper on anxiety using rodent models that didn't use

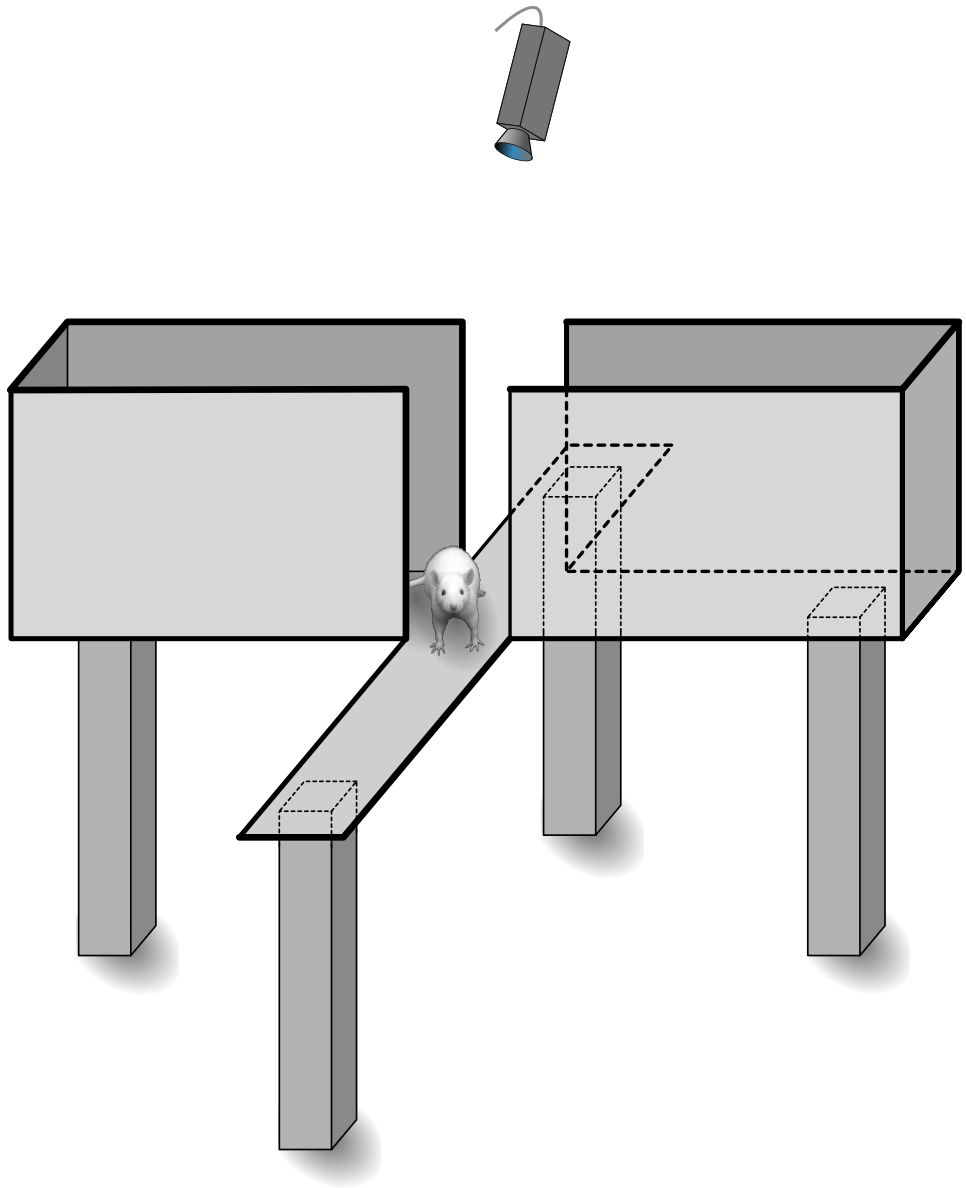

Figure 2: Schematic of an elevated plus maze. Adapted from original work by samuel-john.de. This work is licensed under the Creative Commons Attribution-ShareAlike 3.0 Unported License. To view a copy of this license, visit `http://creativecommons.org/licenses/by-sa/3.0/`.

this test (or a variation of it such as the elevated zero maze). The availability of commercial elevated plus mazes and software programs to automatically measure time spent in each arm contribute to the accessibility of the test, forming a theory-methods package that allows researchers to easily pose questions about what kinds of manipulations change levels of anxiety in mice.

This apparatus is an interesting one to examine for several reasons. First, there is a good amount of information available on the "validity" of the test; that is, whether the test actually measures anxiety like it is supposed to. Lister's initial experiments form the foundation for the validity of the test, but researchers have also continued to analyze different aspects of the maze, such the impact of different maze designs (Hagenbuch, Feldon, & Yee, 2006), different types of behaviors that researchers can measure in the maze (Wall & Messier, 2000), and the test's capacity for detecting anxiety rather than related behaviors like activity levels (Milner & Crabbe, 2008). Second, despite—or perhaps because of—the fact that the elevated plus maze is so widely used, there is still a good deal of discussion about the uses and misuses of this test. The elevated plus maze has a reputation for being a rather finicky test. In the multi-sited study discussed in the introduction, for example, the elevated plus maze was one of the tests that showed the most variation between the laboratories. Researchers have come to various conclusions about why getting stable results out of the test seems to be a challenge. In interviews, some researchers suggested to me that the test was particularly sensitive to environmental variations and that researchers were not taking adequate steps to ensure that their testing room was properly lit and noise free; other researchers argued that the test was being used to ask questions that it was never designed to answer; and others concluded more succinctly that it was simply a bad test. These discussions make the elevated plus maze an appropriate case for looking at how the relationship of particular models to particular disorders gets worked out.

What do researchers think makes this test a good model for human anxiety? The website of *Panlab Harvard Apparatus* (Panlab, 2010), a manufacturer of behavioral test equipment, provides a brief description of the test that summarizes the main technical arguments that researchers use to connect this test to human anxiety:

> The elevated plus maze is a widely used animal model of anxiety that is based on two conflicting innate tendencies: exploring a novel environment and avoiding elevated and open situations constituting situations of predator risk. … When placed into this apparatus, naï¿½ve mice and rats will, by nature, tend to explore the open arms less due to their natural fear of heights and open spaces. In this context, anxiolytics generally increase the time spent exploring the open arms and anxiogenics have the opposite effect, increasing time spent in the closed arms.

There are two main epistemic ladders that are being built in this short description. The first ladder I call the *"pharmacological ladder,"* which is based on the effects that anxiolytic (anxiety-relieving) and anxiogenic (anxiety-inducing) drugs have on the behavior of mice in the maze. The second ladder, which is based on the "innate" tendencies of the mouse to explore some spaces and avoid others, I call the *"ethological" ladder.*

The pharmacological ladder is based on a series of experiments where mice in the maze were given drugs that are known to increase or decrease anxiety in humans. In his original paper on the elevated plus maze, Lister (1987) reported that when drugs such as benzodiazepines (a class of compounds that includes diazepam, better known by its brand name Valium) are administered to mice, they spend more time in the open arms of the maze. Conversely, when Lister gave mice drugs such as caffeine that increase anxiety in humans, it decreased the time the mice spent in the open arms. The behavioral changes that these classes of drugs produce in the maze provide a way of generating links between the mouse's behavior in the maze and human anxiety. Some drugs are known to have anti-anxiety properties in humans based on clinical research, and the pattern of behavioral changes that is observed in mice when they are given these drugs fits the experience of anxiety relief that some humans report when they take the drugs.

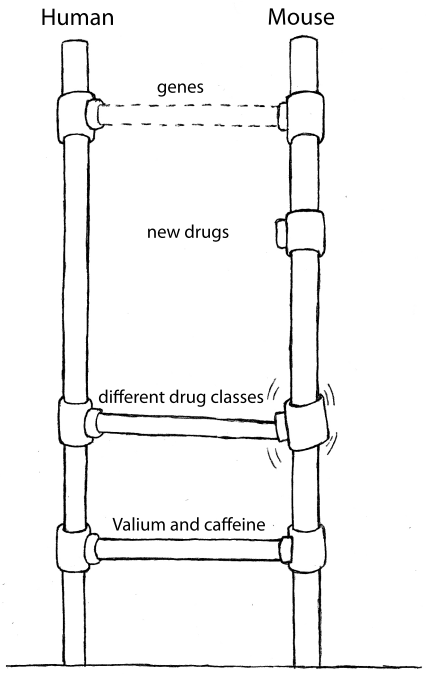

Figure 3: The pharmacological epistemic ladder. Illustration: Nicole Nelson

The observation that the same drugs that increase or decrease anxiety in humans produce corresponding changes in the behavior of mice in the maze forms the base of this epistemic ladder (figure 3). These two aligned pieces of information are linked by the series of experiments performed by Lister to test different anxiety-relieving and inducing drugs such as Valium and caffeine. On their own, the two linked pieces of information may not be that useful. The claim that the elevated plus maze can be used to detect the effect of Valium is a safe claim with a low degree of "externality," but it does not make the test seem terribly useful for doing something other than detecting Valium's effects. But these two linked claims form a foundation from which researchers can build upwards, claiming that this test could potentially be used to find more drugs like Valium, or even more broadly, more drugs not chemically related to Valium that will also relieve anxiety in humans. Going even further, researchers could claim that this test could be used not just to find more drugs that relieve anxiety, but more *things* that cause or relieve anxiety,

like changes in specific genes. Researchers could knock out a particular gene, and then test to see whether changing this gene also changes the mouse's pattern of behavior, just as administering an anxiolytic drug does.

The further up these claims are built, the more open to potential criticisms the ladder becomes. The pharmacological ladder as it appears in figure 3, looks particularly shaky in the upper regions, where the strength (or even the very existence of) the links between the mouse and the human is contested in the animal behavior genetics community. Critics of the elevated plus maze argue, for example, that a key link between the mouse and the human that would support the elevated plus maze is missing, because research with the maze has yet to produce a new drug that has also been found to relieve anxiety in humans (for an example of this critique see Dawson et al., 1995). The lack of a new drug that relieves anxiety in humans leaves the ladder open to possible attacks by actors who question whether the test is actually a good predictor of new classes of anxiolytic drugs. Other researchers argue that the elevated plus maze may not have been used to find any new drugs, but it has been used to produce findings about genes that correspond to genetic studies in humans. For example, a recent study by Cornell University researchers using the elevated plus maze (amongst other tests) identified genes and brain patterns in mice that roughly match genes and brain patterns found in human imaging and linkage studies, which could be presented as evidence of the test's utility (Shmelkov et al., 2010). Further research on other kinds of drugs and the behavioral changes that they induce in the maze has generated a complicated picture at the bottom of the ladder as well. Not all classes of anxiolytics seem to work the same way in the plus maze. While Valium works quite reliably, other drugs that humans report relieve their anxiety like buspirone (an anxiolytic that is chemically unrelated to the benzodiazepines) and fluoxetine (which is part of the class of SSRI drugs commonly prescribed as antidepressants but also has anti-anxiety effects) have shown inconsistent results in the elevated plus maze, sometimes showing no

increases of time spent in the open arms or even slight decreases in time spent in the open arms when they are administered to mice (Moser, 1989; M. T. Silva, Alves, & Santarem, 1999; Kurt, Arik, & Celik, 2000). New drugs that show the same pattern as the drugs that Lister originally tested can be used to strengthen the links between the mouse and the human in the lower rungs of the ladder, but evidence of drugs that do not show the same corresponding pattern can also be used to weaken the ladder by potential critics.

The ethological ladder supporting research with the elevated plus maze draws on an entirely different body of knowledge coming from studies of animal behavior. The ethological argument contends that the test is a good test for anxiety because it provides a simulation of the kinds of situations that mice might encounter in their natural habitat that would be anxiety provoking. Lister notes that the rat version of the maze was developed based on observations from psychologists doing maze experiments in the 1950s who found that if mazes had different sizes of hallways or spaces rats generally preferred to stay in narrow, dark hallways rather than open or brightly light spaces. Likewise, when undrugged mice are tested in the maze, most mice will tend to spend more time in the closed arms. Although Lister uses observations of rodent behavior that took place in a laboratory setting, other researchers often draw from more "ethological" bodies of information about the behavior of mice in their natural habitats. For example, a recently published methodological paper on the elevated plus maze makes a case for the relationship of the test to anxiety by arguing that "mice are prey for many other larger animals, which may underlie their natural tendency to avoid open and, thus, unprotected, spaces (and, to a lesser extent, heights)" (Walf & Frye, 2009, p. 228).

Researchers claim that this test and several other tests of anxiety are based on a conflict between the instinct to explore and the instinct to avoid potentially harmful situations. Mice who are in the maze have an intrinsic motivation to explore the open arms, but how much they do so depends on their assessment of how potentially dangerous this

area of the maze is. Psychologists have termed this predicament an "approach/avoidance conflict," where the instinct to explore a new space is pitted against the instinct to avoid that same space because of its potentially dangerous properties (Wall & Messier, 2000). By tweaking variables that are known to affect the behavior of mice, researchers can make the conditions of the plus maze seem more or less dangerous. Shining a bright light on the open arms of the maze, for example, will make mice even more likely to avoid the open arms than they would under dim light conditions. Researchers argue that this same conflict drives many other tests of anxiety, such as the open field test that involves placing mice in a large, brightly lit arena surrounded by a wall. In this test, researchers measure how much time mice spend at the edges of the maze running along the walls and how much time they spend traversing the brightly lit middle space. The light-dark box, developed by Crawley and Goodwin (1980), also works off of this principle. The apparatus is built like a shoebox with two sides, one of which is covered and dark and the other open and brightly lit. Researchers measure how many times mice emerge from the dark side and how much time they spend in the light.

Researchers also make analogies between the "naturalistic conflict" that mice face in the elevated plus maze and similar kinds of conflict situations that cause anxiety in humans. Jacqueline Crawley, in the aforementioned behavioral testing textbook, provides an example of an anxiety-provoking situation that human researchers might face that is similar to the predicament faced by mice in the maze:

> You want to tell the world about your exciting research results, but have fears about the audience's response when you walk up to the podium to give your talk. A mouse may want to explore a new environment to find food, but may fear venturing out into the open where it is an easy target for predators (Crawley, 2007, p. 230)

Crawley identifies the fear that public speaking generates in humans as the same type of conflict that is faced by the mouse in the maze.

The concept of the approach/avoidance conflict, which is a conflict that researchers assert

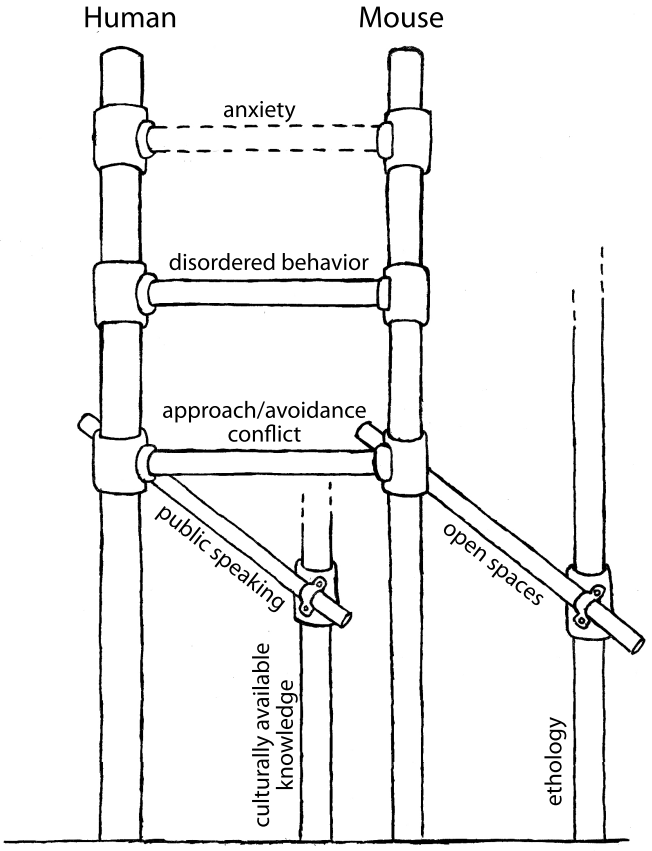

Figure 4: The ethological epistemic ladder. Illustration: Nicole Nelson

both mice and humans share, forms the foundation for this epistemic ladder (figure 4). This concept was developed by German psychologist Kurt Lewin as a theory for describing human behavior, but like many others in psychology after the behaviorist turn, is now used to describe both human and animal action in a deliberately trans-species fashion (Crist, 1999). The claim that the behavior of many different kinds of species can be characterized as an approach/avoidance conflict is already fairly broad in scope, but to create a framework linking the elevated plus maze to human anxiety researchers must build further. From this initial human-mouse link, researchers argue that the test can be used to measure an extreme or disordered version of this conflict behavior that resembles anxiety disorders in humans (such as mice that spend an exceptionally large percentage of their time in the closed arms). Some researchers might even argue that what the elevated plus maze measures is simply "anxiety," rather than just a disordered approach-avoidance conflict. The claim that measurements taken in the elevated plus maze are measurements of "anxiety" is a broad claim that is likely to attract attention and be contested by others in the animal behavior genetics community.

One of the major weaknesses in this ladder is the generalizability versus the specificity of the experiences of the mouse and the human. Some things that will change a mouse's behavior in the maze (like bright lights) are species-specific, but researchers assume that other things (like genes or drugs) will change behavior in both humans and mice. The examples that researchers give to illustrate the approach/avoidance conflict demonstrate this problem. Mice most certainly do not face situations where they are forced to give public speeches, and, as I will explore more below, opinions differ on whether humans would feel more comfortable in the open or the closed arms of the maze. These two pieces of information, which might appear incongruous to outside observers, provide support for this epistemic ladder only if one agrees that the natural history of the species is relevant for identifying what kinds of situations are anxiety producing. Another aspect

of the ethological ladder that generates some skepticism in the field is evidence given for the links between the mouse and the human. The theoretical links that join the mouse and human together in the shared concept of an "approach/avoidance conflict" seems less quantifiable and less "scientific" to some researchers than the links made in the pharmacological ladder.

Both of these ladders touch on different aspects of what behavior geneticists identify as the three types of "validity" that all animal models of a human disorder would ideally have: Face validity (does the animal model look like it represents the human disorder?), construct validity (does the animal model have the same underlying gene mutations or biochemical problems as the human disorder?) and predictive validity (does the animal model predict which treatments will work and won't work in humans?).[2] An ideal model would combine all three of these elements, but researchers argue that few models satisfy all three of these criteria. In developing new models for autism, for example, researchers claim that they can develop tests that have a high degree of "face validity" (that is, tests that resemble the same behaviors seen in autistic patients), but argue that it is more difficult to achieve "construct" or "predictive" validity since the causes of autism are unknown and there are no treatments that reliably address the symptoms (Moy et al., 2004). In many cases, researchers argue that not all of the validity criteria need to be satisfied in order to produce a useful test. The forced swim test, for example, is a test for depression that is widely used despite the fact that many researchers describe the test as one that has little "face validity." The assay involves placing a mouse in a tall glass cylinder full of water and observing how much time mice spend swimming or trying to escape versus how much time they spend just floating in the water. After a period, most mice will stop trying

[2] I have refrained from using these categories so far to avoid introducing unnecessary confusion, because the way in which actors use these terms is not consistent. For example, "predictive validity" might generally be defined as the ability of one measure to reliably predict another. Some researchers use this term to talk about how well one test predicts results on another test (for example Walf & Frye, 2009), while others use the term to talk about how well a test predicts whether or not a drug will be effective in humans (for example Crawley, 2007). I am generally using Crawley's sense of the terms here.

to escape and float in the water, but giving anti-depressant drugs to mice increases the amount of time that they will spend actively swimming (Porsolt, Bertin, & Jalfre, 1977). Although many animal behavior geneticists think that these behaviors have only a passing resemblance to depression in humans, the test is a very successful predictor of drugs that will have anti-depressant effects (Lucki, 1997).

These actors' categories for describing different types of validity can be thought of as different ways of making links between the mouse and the human in the epistemic scaffolds of animal behavior genetics research. The links made in the pharmacological ladder between drugs that change a mouse's behavior in the maze and drugs that alter anxiety in humans demonstrates the predictive validity of the elevated plus maze as a test for human anxiety, because it suggests that the test can distinguish between drugs that will be anxiety-provoking or anxiety-relieving in humans. The ethological ladder's links between the mouse and the human support the face validity of the elevated plus maze, especially if researchers agree that anxiety-provoking situations are species specific. Tests that lack particular forms of validity can be thought of as epistemic scaffolds that are more shaky because they lack the support of particular links that researchers argue have not been achieved. Animal models of human depression lack the support of a "face validity" link between the mouse and the human, because most researchers are not convinced that there is a behavior analogous to human depression symptoms in the mouse's natural repertoire. Similarly, the lack of pharmaceutical treatments for autism makes it difficult for researchers to link the mouse and the human together in ways that demonstrate the predictive validity of animal models for autism.

Each of these ladders' types of links have more or less currency with different practitioners and at different times in the field. Marcus, one of Dennis's postdocs, favored the pharmacological argument for the validity of the test, perhaps because of his background in behavioral pharmacology. Each time we discussed the elevated plus maze, he pref-

aced the conversation with a short description of the pharmacological evidence for the validity of the test. Although pharmacological validation has arguably been considered the most convincing way to make a case for test validity in behavioral research, ethological arguments hold increasingly more sway in the field. In the case of the elevated plus maze, researchers have experimented with introducing new measures that are based in ethological approaches to understanding animal behavior, such as counting instances of "risk assessment behavior" (like peeking around the walls of the closed arms) in addition to measurements of time spent in the open arms (Carobrez & Bertoglio, 2005). The elevated plus maze's ethological relevance is also increasingly used as a marketing point for equipment manufacturers, such as the Harvard Apparatus company who market it as an "ethological test" that is based on "innate behavioral tendencies." (Panlab, 2010)

2.3 Building Up and Breaking Down Epistemic Scaffolds

I have argued so far that the process of creating the epistemic foundations of animal models in behavior genetics can be thought of as a process of building epistemic scaffolds, that are comprised of paired pieces of information about the mouse and the human. The further up researchers build these scaffolds the more broad the claims that they can make using these animal models become, but the claims further up in the scaffolding are also more risky and open to potential attacks. In the case of the elevated plus maze, I identified two different ladders in the epistemic scaffolding of this test, one building on a pharmacological relationship and the other building on an ethological relationship between the mouse and the human. Separating out these two ladders helps to clarify some of the different ways of building relationships within epistemic ladders, but the way that I have described these ladders so far is deliberately static and simplistic. First, the validity of tests like the elevated plus maze is not solely based on the arguments that I have outlined here, but also on other more general arguments about relationships of

similarity of between the mouse and the human. Mouse researchers construct many other epistemic ladders to support the idea of doing research on human disorders using mice, such as those that build upwards from genetic homology, physiological similarity, and evolutionary relationships between the mouse and the human, to name a few. Second, the pharmacological and ethological ladders are not used as separate, distinct arguments in practice, but in combination as part of a global argument for why the elevated plus maze is a good model for human anxiety. In making claims for the utility of behavioral animal models, practitioners are not limited to a single ladder but can use several at once, using the strengths of some ladders to bolster the weaknesses of others. Researchers combine many different epistemic ladders that generate relationships between human disorders and animal models to create a scaffolding to support particular research programs. For example, I attended one training session on the elevated plus maze where I watched the presenter struggle to convince a skeptical audience of new practitioners of the validity of the elevated plus maze as a model for anxiety. She spent several slides outlining the body of evidence for the ethological relationships between the mouse and the human, but when audience members seemed unconvinced she switched to the pharmacological argument to provide additional evidence for the validity of the test. Instances such as these demonstrate how researchers can advance claims upwards by grabbing onto the rungs of other ladders, moving fluidly through a landscape of existing arguments about the relationship between animal models and human disorders.

This section pays deeper attention to how researchers break apart and reconfigure specific aspects of the epistemic scaffolding of the elevated plus maze. Even though the epistemic foundations of tests have a tendency to become black boxed over time, researchers can reveal and modify a test's foundations by publishing new research on test validity, reviewing each other's papers, teaching graduate students about the origins and evidence for particular models, and repeating particular descriptions of these models in

everyday interactions in the lab. In most of the examples that I collected in my field notes and interviews, researchers were trying to break down specific aspects of the epistemic scaffolding for the elevated plus maze rather than building it up. Many of the behavior geneticists that I interacted with thought that the utility of this particular test was often overstated, with too much emphasis on its capacity to capture "anxiety" and too little emphasis on its limitations. In this section, I explore some of the specific links and pieces of information in the epistemic scaffolding of the elevated plus maze that the researchers I interviewed questioned, and I outline some of the general techniques that researchers use to modify epistemic scaffolds.

One of the ways that researchers in the Smith lab regulate the distance between the animal and the human is by paying close attention to the language that they and their colleagues use to describe what is going on in tests like the elevated plus maze. Researchers employ careful language to indicate that the behaviors of mice in the maze are mouse behaviors, and not the behaviors or feelings of miniature humans. I described one such incident in my field notes, where I was corrected on my description of mouse behavior by Marcus while watching him score video tapes from elevated plus maze experiments over lunch. Marcus started, as he typically did, by giving me the pharmacological argument for the validity of the elevated plus maze:

> After he had finished his blurb and I'd asked some questions, I just sat back and watched a few rounds while I was eating my salad. At one point in watching the video I commented, "That guy really likes the open arms," because the mouse that we were watching at the time seemed to be spending more time there. There was silence for a minute, and then (as though my comment had just registered to him and he was slow to respond because he was thinking about other stuff), Marcus said, "Don't say 'like.' " It took me another minute to figure out what he meant here, and then it dawned on me that he was chastising me for "anthropomorphizing" the mouse. I asked him about this, and he said that he had been chastised before by other people for saying things like what I had just said. He said that you should never say things such as "the mouse likes the open arms" or "the mouse is less anxious," you should say things like "the mouse spends a higher percentage of time in the open arms" or "the mouse shows less anxiety-like behavior."

Marcus thought that my comment that the mouse "liked" the open arms was an inappropriately anthropomorphic representation of what was going on in the maze. The phrases that he suggested instead—"the mouse spends a higher percentage of time in the open arms" and "the mouse shows less anxiety-like behavior"—refocus attention away from what the mouse might be feeling and towards what experimenters can observe.

Researchers in the Smith laboratory often refer to the elevated plus maze, as Marcus did, as a test for "anxiety-like behavior" rather than as a test for anxiety. This somewhat cumbersome construction inserts some distance between the animal test and the human behavior and lowers the epistemic scaffolding test to a less risky claim: It denotes that what the mouse is doing in the maze is aligned with human anxiety, but that it may or may not be the same thing as human anxiety. The animal behavior geneticist's vocabulary is replete with terms like these that modify and restrict the meaning of what the mouse is doing in behavioral tests: "Anxiety-like behavior," "depression-related behavior," and behaviors "relevant to schizophrenia" are just a few more examples of the kind of hedging constructions that the Smith lab researchers use. At Western, researchers use these terms of art in both formal and informal situations until they become almost second nature—in class presentations, in poster sessions, and in conversations with each other. Although they are more careful about ensuring that they use this cautious language in certain venues, they also use it even in the more private settings of the laboratory. The Smith laboratory members, for example, ironically applied these phrases to themselves and others in the lab as a source of humor. Sharon might comment that "Delores is really stressed about the chamber study," and then joke that she should say instead that "Delores is showing stress-like behaviors." This type of joke was typical of Sharon's style as a laboratory manager, where even during humorous moments she still retained her self-described tendencies to be "picky" and "perfectionistic."

Careful constructions such as "anxiety-like behavior" push back on moves to build up

toward the statement that both mice in the maze and humans experience "anxiety," such as the moves that I identified in the description of the ethological ladder. These phrases also create distance between animal models and human disorders by inserting some generalized instability into the scaffolds connecting the mouse and the human. Calling these behaviors "anxiety-like" rather than "anxious" suggests that there are some disconnects between models and human states without necessarily identifying any particular weak links.

Other linguistics moves attack particular links in epistemic ladders. In the following example from an interview with Ethan, a graduate student in a psychiatric genetics program, he corrects my description of the tests that we are talking about which I have referred to as "tests for depression" and replaces it with the term "tests for anti-depressants":

> *NN:* So most of the other tests out there for depression all have that time lag problem where they all respond to acute effects?
>
> *Ethan:* Yeah, yeah.
>
> *NN:* Like the forced swim test?
>
> *Ethan:* I mean, tests for anti-depressants, right? I've been trying to be a little bit more vigilant about separating what's a model of depression versus what's a test of anti-depressant efficacy, right? Because they're not necessarily the same thing.

Here, Ethan interrupts the conversation to (nicely) correct me on the way that I have named this test. He thinks that the tests we have been discussing here should be described as tests that reveal the effects of drugs that act as anti-depressants in humans. He notes that there is a difference between saying that the tests he uses identify drug effects and saying that the tests model "depression." His reformulation shortens the pharmacological ladder, suggesting that the forced swim test can be used to model responsiveness to drugs but not to model depression in a more general sense. Sharon often made a similar argument in conversations about the elevated plus maze. She thought that the test worked well for some purposes, such as an initial screen for drug discovery, but that when researchers used the test as a test for "anxiety" they were likely to run into more serious interpretational problems.

Other researchers thought that calling tests like the elevated plus maze "models for anxiety" was appropriate, but only if researchers acknowledged these tests probably modeled only one component of anxiety and not the entire disorder. They also avoided referring to the plus maze as a "test for anxiety," because they worried that this terminology encouraged animal researchers to mistake the one aspect of anxiety that the test measured for the global concept of anxiety. Scott, a behavior geneticist who has had extensive experience with the elevated plus maze, articulated this concern about its interpretation when I interviewed him about his research. He elaborates:

> Are you measuring anxiety, or are you tapping into some undefined piece of an anxiety dimension that's hard enough to diagnose in a human? Let's not pretend that this test is a measure of anxiety behavior. No, it's a measure of behavior that's probably somehow related, but let's not start substituting the label anxiety for this behavior.

He suggests that the test is "probably somehow related" to anxiety, but what it measures is likely only one facet of the much larger phenomenon that we call anxiety in humans. Since it is difficult enough to define anxiety in humans, he thinks that it is unreasonable to suggest that researchers have captured the full complexity of anxiety in a mouse test. Scott is also pushing down on the epistemic scaffolding of the elevated plus maze, arguing that researchers should make the more circumscribed claim that their animal models may capture some part of anxiety, but not everything that might be identified as anxiety in humans.

Scott's comments are primarily focused on the uncertainty on the human side of the epistemic scaffolding about what exactly constitutes anxiety, but Dennis argues that there is also uncertainty on the mouse side about what constitutes "anxiety-like behavior." He points to a set of experiments that he published where he attempted to group tests for motor coordination into sub-categories based on more specific features of motor coordination, like balance or the gripping strength of a mouse's paw. Even with his knowledge of animal behavior, he says that in this case "anthropomorphizing didn't help [him] at all," because

he turned out to be quite wrong about the relationship of tests that he thought were very similar. He started with a test called the "grip strength test," where mice grab with their forepaws onto a small bar attached to a force meter that measures how hard they can pull:

> So, you do that, you measure that, and alcohol makes them weaker, okay? And then we add another task where you have a window screen in a frame, and you take the mouse and inject it with alcohol, put it on the window screen and rotate it so it's horizontal like this, and wait and see how long it takes before it falls off.[3] With saline injection they can run around up there forever, but alcohol makes them fall off. No relationship whatsoever between those two tasks. So our ability to guess what each of those tasks was measuring was about nil.

Dennis hypothesized that both of these tests measured similar aspects of motor coordination that had to do with a mouse's ability to hold onto the bar on the grip strength apparatus or the mesh on the window screen. But his experiments revealed no relationship between these two tests: Strains of mice that held onto the window screen for a long time while intoxicated didn't necessarily do well on the grip strength, and vice versa.

This story weakens the epistemic scaffolds of animal behavior genetics research because it suggests that researchers' abilities to identify behaviors in mice that can be linked to similar behaviors in humans might not be as accurate as they think. Dennis initially thought that both tests might be measuring something like "grip strength," but found that the two tests seemed to be measuring different mouse abilities. He suggests that there might be similar problems with animal models for anxiety, and that not all tests that researchers describe as involving an "approach-avoidance conflict" are necessarily measuring the same thing. For this reason, it has become more common in the animal behavior genetics field to use multiple tests to measure the same behavior to ensure that the results are not idiosyncratic to a particular test. While ten years ago papers on anxiety

[3] In this test, the researcher places a mouse in the middle of a square frame covered with fine wire mesh that resembles a window screen. At the beginning of the test the window screen is horizontal, and researchers gradually tilt the screen (either by hand or using a robotic arm) so that the surface the mouse is standing on is increasingly slanted, then vertical, then tipped so that the mouse is on the underside of the screen. Researchers measure how long the mouse can hang onto the tilting screen before it loses its grip and falls off.

could be published in high-profile venues with data from just the elevated plus maze, many research groups now use at least three separate tests for anxiety, such as the elevated plus maze, the open field test, and the light-dark box.

Finally, some researchers identify entire ladders, not just links or alignments within the ladders, as inappropriate. Amy, a veterinarian who works for a major commercial supplier of inbred mice, was skeptical about any kind of ethological argument for the validity of the elevated plus maze. I met her at the *Measuring Behavior* conference after she had attended a one day workshop outlining the available range of mouse tests for stress and anxiety, and she told me that the presenter's comparison of approach-avoidance conflicts in mice and anxiety in humans, to her, simply did not ring true. She thought that if anything, the test looked like it measured something more akin to "bravado" rather than "anxiety," and that claiming that the test measured anxiety was too anthropomorphic. She found the pharmacological evidence for the test more convincing, but was still skeptical about whether this experimental data could provide any information about whether the mice themselves were anxious:

> I think that for the [tests] where it has been pharmacologically proven that anxiolytics decrease that behavior, I think there's an argument for it. But I'm still not going to say that the mouse is anxious and we're alleviating it. All I'm saying is that whatever pathways that alleviates anxiety in us exist in the mouse and we're alleviating that as well. Now whether [the mice] interpret that as anxiety or whatever else, I don't know if we can make that leap, scientifically.

For her, the pharmacological ladder provides convincing evidence that there is a similarity between mice and humans, but she characterizes this as a biological similarity, not an experiential one. The fact that human drugs change behavior in the mouse might mean that they have the same receptors where the drugs might act, but it does not mean that the mouse also has a mental state that we can equate to that of humans. The elevated plus maze measures some kind of change in the mouse, but she is unwilling to say that what it measures is "anxiety."

2.4 Don't Anthropmorphize! Epistemic Scaffolds as Shared Frameworks

So far I have demonstrated some of the different ways in which researchers build up and break down epistemic scaffolds in their practice, from publishing experiments that question the ability of researchers to link the human and the mouse, to reminding each other to use terms like "tests for anti-depressants" rather than "models for depression." These techniques for modifying the epistemic scaffolding that supports models like the elevated plus maze are part of a process of calibrating the type and strength of claims that can be made with that particular test. The finely-tuned epistemic scaffolding that result from this process of calibration acts as a framework for understanding different aspects of the work of animal modeling, such as what animal models measure and how much information they can provide about human disorders.

This section revisits the contradiction that I started with in the beginning of the chapter; namely, that animal behavior geneticists caution both outsiders and each other not to "anthropomorphize" while simultaneously engaging in practices that appear to the outside observer to be quite anthropomorphic. I argue that this caution and others are ways of *indicating* shared frameworks for understanding the practice of animal modeling in behavior genetics (Goffman, 1974). Eileen Crist (1999) argues that there is no hard and fast definition of what counts as anthropomorphism; rather, "its meaning is tied almost strictly to its aspersive connotations, for it suggests a manner of representation entailing the figurative, erroneous, or naive attribution of human experiences to animals" (p. 7). Taking Crist's cue, I suggest that there is no clear delineation between activities in the Smith laboratory that are anthropomorphic and those that are not, but that "anthropomorphizing" is a term that actors use to mark out what they consider to be *inappropriate* links between the mouse and the human.

Some important features of this shared framework are by now visible. First, the types

of information that researchers align and the links that they make largely exclude the animal mind as an appropriate topic for investigation, a position that has a long tradition in behavioral psychology (Crist, 1999). Humans' subjective experiences (such as self-reports of the effect of certain drugs) can figure into the epistemic scaffolding of animal behavior genetics, but the corresponding information on the mouse side focuses on mouse behavior, not subjective experience. While a few researchers might want to climb up to the shaky regions of the epistemic scaffolding to claim that both mice and humans experience anxiety, researchers in the Smith laboratory agree that animal models are better used to explore behavioral or biological similarities with humans. Second, the Smith laboratory researchers agree that animals and humans are similar in some ways but not exactly the same, and that animal models never capture all of the features of a human behavioral disorder. They point out that there are species-specific features of behavioral disorders such as the differences in what mice and humans find anxiety provoking, and that the correspondence between things like the drugs that humans report relieve anxiety and drugs that change mouse behavior in the elevated plus maze is rarely perfect. Within this shared framework, researchers differ considerably on what they think is similar. While some argue that the strongest alignments between the mouse and the human are to be found in common patterns of behavior, others argues that biological features like shared receptors constitute much stronger links.

Engaging in certain kinds of "anthropomorphic" thinking and speaking in the laboratory is a way for researchers to signal to each other that they share a similar understanding of what the epistemic scaffolding connecting the mouse and the human looks like. As I mentioned previously, researchers engage in many activities that look, at least to an outside observer, anthropomorphic. Consider for example this exchange with Sharon, the Smith laboratory manager, who is trying to explain to me why testing mice who are withdrawing from alcohol in the elevated plus maze presents problems:

> *Sharon:* When you're looking at say withdrawal-induced anxiety, then you have a problem. Because during withdrawal … imagine you've got a hangover, what are you going to do?
>
> *NN:* I'd be sitting the middle.
>
> *Sharon:* You'd be sitting, exactly.
>
> *NN:* I would be hanging out and waiting until my time was up and I could go back to my cage!
>
> [laughter]
>
> *Sharon:* Exactly! And so mice that are withdrawing from alcohol don't move very much. So how do you interpret their behavior on that maze? You can't. So you then don't know if it's behavioral, they're not motivated, you don't know if it's malaise, they just don't feel well. Like every time they lift their head the room goes like that, you know? [acting out a spinning motion]. There's no way to interpret those data.

To explain the issue to me—that mice who are withdrawing don't move enough to get reliable information on how much time they spend in the closed arms versus the open arms—she invites me to think about what I would do if I had a hangover as a way of reasoning through the problem. I respond by placing myself in the position of the mouse in the maze, which might appear to be a clear instance of anthropomorphizing. Sharon does not correct me, however; she joins me in this activity. She also embodies the withdrawing mouse, acting out a mouse whose head is spinning from a rather nasty hangover.

One of the features that distinguishes between appropriate and inappropriate uses of this mode is that scientists (and sociologists) must imagine and talk about themselves *as a mouse*, not just as a small human in a maze. Researchers must maintain an appropriate distance between mouse and human, and not fully transpose one onto the other in the way that they talk about mice in the lab. In the example above, it is clear that I am at least trying to speak like a mouse who wants to return to her cage, and the actions that I describe are deemed by Sharon to be compatible with mouse-like behavior. Speaking from this mouse-like perspective requires a good deal of practice, because it is easy to slip back into speaking from the perspective of a "furry, one-ounce human" during these moments, as one veterinarian put it to me. On other occasions when I adopted the same

mode and began to describe what I might be seeing off of the end of the maze on the floor or an adjacent wall, I was quickly corrected: Mice have relatively poor vision and rely very little on their sense of sight, researchers told me, indicating that describing what you-as-mouse might see from the open arm of the maze is not compatible with the framework of species-specific modeling.

Designating some activities and not others as "anthropomorphic" circumscribes and reinforces the boundaries of this shared framework, and it also indicates what (or who) falls outside of it. Cautions that researchers should not anthropomorphize can also be understood as part of a professional struggle for control over who gets to make a claim about anxiety using the elevated plus maze. The tensions that I alluded to earlier between those who consider themselves experts in animal behavior and those that they identify as part of the neighboring field of molecular biology are especially relevant here. Several researchers at Western argued to me that it was almost too easy to gain access to behavioral tests like the elevated plus maze. Unlike other kinds of more expensive and cumbersome lab equipment, a researcher could easily borrow a maze from a neighboring lab from down the hall, or buy a commercially-produced maze and a video camera for a few thousand dollars. The test is so easy to run that it is almost impossible to not get data: If you place a mouse in the maze, it will spend some percentage of time in the closed arms and some percentage of time in the open arms. At the end of this experiment, you can analyze this data and make some claim about whether your treatment increased or decreased anxiety levels in the mice. This kind of scenario where a researcher might casually buy a maze, test some mice, and conclude that a mouse is more anxious because it spends more time in the open arms is immensely problematic to the researchers at Western. Graduate students at Western complained that molecular biologists were often "not thoughtful" about interpreting the meaning of these tests, and could "pollute the literature" with bad data if they didn't properly control for factors that might increase or decrease mouse

anxiety (such as loud music or changes in the light levels in the test room) or carefully consider the meaning of the data from their tests.

Anthony, an animal behavior geneticist from Canada, suggested that there should be a kind of disciplinary division of labor, where behaviorists make and validate tests that the molecular biologists can then use once they are appropriately vetted and packaged. I asked him:

> *NN:* If you put people from psychology and molecular biology backgrounds together and tried to get them to figure out what the ideal protocol would look like, do you think they'd have different approaches to trying to do that?
>
> *Anthony:* Well, I wouldn't let the molecular biologists have anything to say about it!
>
> [laughter]
>
> *Anthony:* Because what do they know about plus mazes? I mean, this is something where the behavioral psychologists have to get together and provide good, validated tools that the molecular people can then use.

Anthony emphasizes that the job of deciding which tests adequately reflect human behavioral disorders is the job of behavioral psychologists, who can use their knowledge of mouse behavior to better align the mouse and the human. He thought that the elevated plus maze, despite its widespread use, was not a test that was ready for wide circulation because it still had substantial methodological and interpretational issues to be worked out. Thus, the warning "Don't anthropomorphize!" could be seen as an assertion from the animal behaviorists that a competent practitioner needs to have an understanding of mouse behavior, not just a maze and a video camera. Researchers' assertions that certain activities or ways of talking about the mouse are inappropriately anthropomorphic centers the discussion around knowledge of mouse behavior and the species-specific elements in the epistemic scaffolds of tests like the elevated plus maze. Talking from the perspective of a mouse also requires a good deal of knowledge about the behaviors and sensory words of mice, and emphasizes both one's knowledge of mouse behavior and commitment to taking it seriously as an important part of animal behavior genetics practice.

Finally, creating and enforcing a shared framework for animal behavior genetics work is also part of a larger problem of how to shape claims in order to maintain the credibility of the field. Animal behavior geneticists express concerns that members of the general public are also likely to think that they understand behavioral testing and over-interpret the significance of behavior genetics research. Anecdotally, my experiences talking about the elevated plus maze suggest that even people who are quite unfamiliar with animal behavior genetics research often feel knowledgeable enough to talk about the test after only a brief introduction to it. When I have presented papers about the elevated plus maze at conferences, I found that even in an audience made up of social scientists and historians some commenters challenged the validity of the test by proposing alternative explanations of what the mouse might be doing in the maze, or by offering descriptions of what they would do if they were in the maze. At a conference where I gave a version of this chapter, a member of the audience commented that she didn't find this test to be a plausible model for human anxiety because a human in the maze would do the opposite of what the mouse does—go out into the open arm and have a good look around. Another audience member approached me afterwards to say that she thought the first comment was absolutely wrong, and that she would stay in the closed arms of the maze. Others have suggested that perhaps the mouse is not afraid of open spaces but afraid of heights, and that perhaps the floors should not be made of clear plastic. This test seems to be particularly readable to a non-specialist audience, despite the complexities involved in its production. Rebecca Lemov (2006) argues that similar maze experiments with rodents conducted in the 1950s and 1960s caught the public imagination because they were easy to relate to human situations. She argues that they provided miniature "dioramas of real life" (p. 5) that resembled the human world but had added authority because they took place within the controlled setting of the laboratory. The conflict presented in the elevated plus maze—to explore open space or stay in an area of safety—is one that draws on commonsense ideas

of the world in ways that allows people to picture themselves in this world made for mice.[4]

Careful constructions such as "anxiety-like behavior" provide a linguistic reminder that is also aimed at the public about the distance between research with animal models and the ultimate goal of understanding human conditions. Like the "turn towards complexity" narrative that Arribas-Ayllon et al. (2010) identify, cautious language helps to manage expectations about what the field is doing and what the public can expect it to produce in the future. In her textbook on mouse models, Crawley (2007) explicitly makes this link between cautious language and maintaining the credibility of the field. She says:

> Tread softly when approaching a mouse model of a human psychiatric disease. Investigators have no insight into whether a mouse feels "anxious" or "depressed" … Members of our laboratory are taught from day one to use cautious terminology such as "anxiety like," "depression related" and "relevant to schizophrenia." The credibility of our field depends on avoiding the impression that it is possible to create a comprehensive mouse model of a human mental illness (p. 261–62).

In an interview, I asked her to elaborate on whether she was thinking about maintaining credibility in the eyes of other scientists, or about credibility with other groups like funding agencies and the public. She replied:

> We're never going to have all the components of a human disease in a mouse model. We break it down as much as we can, we get as many of the elements that are the critical, core diagnostic symptoms in as we can. So we have to be modest and understand the limitations that mice are never going to recapitulate the entire human syndrome, and we have to state that very clearly in our papers that we're studying anxiety-like behaviors, or an antidepressant-responsive task, or social deficits relevant to autism, not that we have an autistic mouse. We try never to say anything that sounds like we have a blank mouse. People over-interpret very dramatically. The popular literature, I guess, and newspaper and magazines, you know, that's the headline. And it's just not true.

Behavior geneticists perceive themselves to be caught in a web of conflicting pressures when it comes to how much they can say about their models, many of which I have alluded to already. Failures to replicate findings from previous studies have made animal

[4] See also Brannigan, 2004 for a similar argument about social psychology experiments.

behavior genetics generally cautious about making strong arguments for links between a gene and a behavioral trait, but as I will explore in more detail in later chapters, there are many other tensions that shape the culture of claims-making in animal behavior genetics. Many of today's senior behavior geneticists, who began their careers during the often vicious debates about intelligence and heredity in the 1960s, are also wary of the public implications of overly determinist statements about biology and psychiatric traits. On the other hand, researchers want to demonstrate to funding agencies, to other researchers and to the public that animal models do have an important role to play in understanding and treating behavioral disorders, even if they can never fully recreate the complexity of the human experience. In particular, animal behavior geneticists perceive themselves to be in a longstanding war with animal rights groups about the utility of animal models, and feel compelled to counter forceful arguments from activist groups that animal models are not useful in understanding human health at all. Making claims in animal behavior genetics is not just about the representativeness of models of human disorders, it is also about managing these tensions around how much to claim about the relationship between genes and behavior.

2.5 Conclusion

This chapter has analyzed the ways in which animal behavior geneticists build the relationship between their models and the human disorders that they study, looking at how the animal and the human are drawn together and pushed apart by researchers. To better understand how animal models model, I developed the concept of an *epistemic scaffolding* supporting research with animal models that involves both a horizontal process of *linking* specific information from the animal and the human and a vertical process of building increasingly broader and riskier claims about their relationship. Examining the arguments that researchers make for the validity of the elevated plus maze shows how researchers

use particular pieces of evidence from both the mouse and the human as the basis for building up ladders that support this test as a tool for studying human anxiety disorders. The elevated plus maze has several differently configured ladders, such as one based on pharmacological evidence and another based on ethological evidence, that are combined into a single epistemic scaffolding that supports the test.

This scaffolding can be built up and broken down by practitioners in the animal behavior genetics community. The researchers I interviewed frequently argued that others within the field overemphasized the relationship of the elevated plus maze to human anxiety. These types of criticism are not only about the strength of the available evidence for the validity of the test, but they may also arise from concerns about claims that other researchers are making using the test, or from struggles to assert behaviorists' professional authority in the broader scientific community. Those who were critical of the elevated plus maze employed several techniques for attacking specific links or the overall height of the test's epistemic scaffolding. Calling the elevated plus maze a test for "anxiety-like behavior" rather than a test for "anxiety" introduces some uncertainty into the scaffolding linking the test to the human condition. Other arguments, such as the suggestion that the test should be thought of as a test for anti-anxiety drug effects rather than a test for anxiety, attack specific links in the ladder that researchers think are unwarranted. At the same time, these moves also increase the stability of the epistemic scaffolds by reducing them to a safer height, making animal behavior genetics less vulnerable to criticisms that they are overselling the capabilities of their models.

Finally, I argued that a general agreement around the appropriate configuration of the epistemic scaffolding connecting the mouse and the human constitutes a *shared framework* for understanding the practice of animal modeling at Western. By marking out some activities but not others as inappropriately "anthropomorphic," researchers reinforce their shared framework and delineate which practices and practitioners fall outside of

it. Researchers at Western emphasize that interpreting tests like the elevated plus maze requires a familiarity with mouse behavior and an understanding of the limitations of these tests, knowledge that some believe the public or practitioners from the neighboring field of molecular biology may lack.

My goal in this chapter was not to make my own evaluation of the strength of the relationship between tests like the elevated plus maze and human disorders like anxiety, but to show systematically how researchers themselves make and manage these associations. Warnings such as "don't anthropomorphize!" reflect tensions that are specific to the field of behavior genetics and may not necessarily be evident in other fields where animal models are used. While cancer researchers might see their work as just as complex and multi-faceted as the practice of modeling behaviors, it is probably unlikely that researchers who use animal cells to model human tumors caution new practitioners against the dangers of imbuing their mice with too many human-like characteristics. The story that I have told about the elevated plus maze might also not be the same for all behavioral models. Tests like the conditioned place preference test described in the previous chapter might have less disagreement around their interpretation because they are packaged with more specific theoretical concepts, such as psychological theories about drugs as "reinforcing" or "rewarding," and about the ways in which humans and animals develop associations between environmental "cues" and the experience of being on a drug.

Although it is advantageous for researchers to make strong claims about the capacity of animal models to reflect human disorders, the space between animals and humans is not a space that researchers want to completely collapse, as evidenced by the efforts of Western researchers to deliberately introduce uncertainty into the relationship between mouse and human in their own models. There are many instances in which researchers might strengthen their professional position by reducing the scope of their claims. Emphasizing uncertainties in the relationship of animal models to human disorders helps animal

behavior geneticists retain professional credibility and control over the tools of their trade that are increasingly being taken up in other disciplines. By making, managing, and breaking specific links in epistemic scaffolds, animal behavior geneticists are carefully calibrating the strength of the claims that they make about the relationship of animal models to human disorders, simultaneously drawing together and pushing apart the human and the mouse.

3 Reducing Humans to Genes? The Metaphorical Entailments of Animal Behavior Genetics Models

Taped to the wall next to the door of a laboratory at Western University is a copy of a photograph showing a middle-aged man on a street corner. The man is sitting on an overturned milk crate, with an empty McDonald's cup placed in front of him. His heavy blue jacket and black pants blend in with the grey street scene, and his face is turned away from the camera so that only the back of his baseball cap and his ponytail are visible. In the center of the photograph is a cardboard sign written in black marker that the man is holding out. The sign says, "Need cash for alcohol research." Underneath the photograph, someone has drawn an arrow pointing to the man in the picture and added a handwritten note that reads, "Dr. A when he had long hair."

Who is the anonymous man in the picture? A witty panhandler? An alcoholic? Or, as someone at Western has jokingly suggested, a fellow alcohol researcher in need of funding? The humorous photograph plays on culturally available ideas of who the alcoholic is: An older man, alone, without a job, and possibly homeless, whose day is organized around getting alcohol or money for more alcohol. The image may also conjure up ideas of how the alcoholic came to be this way, and how society should respond. Some might see a man who has chosen a destructive lifestyle that has isolated him from friends and family, and others might see a man who is suffering from the disease of alcoholism. Perhaps his present situation was brought on by a genetic predisposition to alcoholism, or by stressful

life events that pushed him towards heavy drinking, or by choice. Maybe the man in the image needs to take responsibility for his drinking, or admit that he has no control over his consumption of alcohol. Maybe new medical interventions or public health programs are needed to help him stop drinking, or to keep others like him from drinking in the first place.

Commentators from many disciplines, both inside and outside of the sciences, have been critical of how behavior geneticists' search for genetic contributors to alcoholism shapes this image of the drinker. Sociologists have argued that genetic research flattens a complicated picture of disease causation, focusing attention on genes as the primary causal factors for disorders such as alcoholism and pushing sociological contributors to health and illness out of the frame (Conrad & Schneider, 1980; Conrad & Weinberg, 1996; Conrad, 1999a; Nelkin & Lindee, 1995; Room, 1974, 1983; J. W. Schneider, 1978). Researchers have coined the term "geneticization" to describe the growing tendency for scientists, medical professionals, and laypeople to attribute human differences to genetics (Lippman, 1991), with potentially detrimental consequences (Lewontin et al., 1984; Duster, 1990). Others argue that genetic research on alcoholism legitimates the idea that it is a "disease" rather than a social problem, an assertion that they argue lacks factual support and encourages drinkers to adopt a fatalistic attitude towards their drinking behaviors and prospects for changing them (Szasz, 1961; Fingarette, 1988).

This chapter looks at how representations of the human are articulated through the process of mouse research. As Eileen Crist (1999) has comprehensively demonstrated in her research on language in the behavioral sciences, the terms that scientists use to describe animal behavior have substantial imagistic and conceptual effects on the way that readers understand animals' actions and the relationship between animal and human worlds. The experimental systems and associated epistemic scaffolds that researchers construct to study human behavior using animals also have consequences for how researchers concep-

tualize human behavior. In the previous chapter, I argued that particular arrangements of the epistemic scaffolding supporting animal models constitute shared frameworks for understanding the work of animal behavior genetics. Looking in depth at a particular behavioral test, I explored how practitioners adjust the strength and configuration of the scaffolding supporting that test, declaring some links between the mouse and the human to be sound and others to be unsupportable. Social practices in the research community, such as using cautionary language, reinforce the idea that the mouse and the human are linked together in specific, circumscribed ways in the work of animal modeling. In this chapter, I examine the images of the human that are embedded in these tests and epistemic scaffolds, asking: What features of human disorders do researchers emphasize or exclude from their animal models, and what do they offer as their reasons for doing so? How are the analogies between the mouse and the human extended or curtailed in the practice of animal behavior genetics research? What happens to researchers' visions of the human drinker after they are run through animal experimental systems? Is the process of animal modeling "reductionist," and if so, what does it reduce alcoholism or the human to?

I draw from interviews and ethnographic material at two sites where the Smith laboratory researchers interact with other scientists around common research objects, such as shared protocols or inbred strains. The majority of the chapter follows the development of a new mouse model for binge drinking behavior in an interdisciplinary consortium group that I will refer to as the Alcohol Research Consortium (ARC). Developing a mouse model for binge drinking is a particularly appropriate site to examine how researchers negotiate what features of human drinking are essential or practical to represent in animal models, because getting mice—who are not naturally inclined drink alcohol—to consume it in large quantities requires substantial experimental manipulation. I look at how the members of the ARC articulate the human concept of "binge drinking" through the recalcitrant body of the mouse, and how the links that they make between human binge

drinkers and their experimental mice offer resources for reflecting on human drinking. In the final section of the chapter, I introduce a second example to explore how this process of articulating the human through animal research operates in a project where researchers are modeling the human in a general sense, rather than modeling a specific kind of human behavior. I describe the *Mouse Phenome Database* (MPD) project, an interdisciplinary initiative to collect phenotype data and protocol information from mouse researchers, and examine how the coordinators and users of the database conceptualize the laboratory mouse's environment and how it relates to human environments and health.

I suggest that the epistemic scaffolds for these models and databases can be thought of as extended, bi-directional metaphors that researchers can use to highlight particular features of either the human or the mouse. By examining some of the various ways that researchers use these models as analogies for the human situation, I argue that human disorders are not straightforwardly "reduced" to genes by the animal models researchers develop to study the genetics of human alcoholism. In particular, I focus on the way that researchers use animal models as resources for talking about not only genetic predispositions to drinking, but also environmental factors that control human drinking. While researchers are aiming to produce information about genetic factors that make some humans more susceptible to addiction than others, they also use the process of setting up experiments and developing new models to highlight environmental effects on behavior.

3.1 Reductionism and Metaphor in Animal Behavior Genetics

One of the criticisms frequently leveled at behavior genetics research is that it encourages "genetically reductionist" thinking that emphasizes the importance of genes while de-emphasizing other things that also shape and give meaning to human behavior. Behavior geneticists' claims to be able to assess the heritability of human traits such as intelligence or aggression—or even to identify specific genes for these traits—has inspired spirited

criticisms and objections from commentators both inside and outside of the biological sciences. This section reviews some of these critiques of behavior genetics research, with a focus on different ways that analysts have conceptualized the production of genetically reductionist ideas in the field. While some suggest that behavior genetics methodologies (and perhaps behavior geneticists themselves) are inherently reductionist, others describe the field's practitioners as aware of some of the problems of reducing complex behaviors to genetic explanations but constrained by practical limitations and existing cultural metaphors for articulating concepts of gene action. I suggest that the epistemic scaffolding of animal behavior genetics models can also be thought of as metaphors that provide resources for scientists to highlight or mask particular features of the human, offering not just constraints but also new possibilities for understanding human drinking.

In his book *Genetics and Reductionism*, philosopher Sahotra Sarkar (1998) outlines the concept of reductionism and how it operates in genetics. He defines genetic reductionism as the process of explaining one set of phenomena in terms of genetic phenomena, or "the thesis that all phenotypic phenomena can (always) be reduced to facts at the genotypic level" (p. 10). He distinguishes this from the more general concept of "physical reductionism" (the idea that all biological phenomena have a physical basis) and "genetic determinism" (the idea that having a particular genetic allele ensures that an individual will also possess a particular trait). Sarkar, like many outside observers of the behavior genetics field, concludes that reductionist approaches in fields such as plant breeding or molecular biology have some merits, but he remains unconvinced by the application of these approaches in behavior genetics. Biologists and even practicing behavior geneticists have also been critical of the field's attempts to reduce human behaviors to genetic and environmental components through methods such as heritability studies that assess the proportion of a trait that can be attributed to genetics in a population (see for example Lewontin, 1974/2006; S. Rose, 1997; Wahlsten, 1990). Commentators have further main-

tained that the tendency of behavior genetics to reduce human traits to genetics deserves particular scrutiny because these claims have the potential to negatively impact public policies and perceptions of human behavior (Alper & Beckwith, 1993; Duster, 1990; Kevles, 1992; Nelkin & Lindee, 1995). Steven Rose (1997) argues that one of the immediate social consequences of reductionist thinking in biology is that "attention and funding is diverted from the social to the molecular. If the streets of Moscow are full of vodka-soaked drunks … then [reductionist thinking] demands the funding of research into the genetics and biochemistry of alcoholism" (p. 297). In their book *Not in Our Genes*, Lewontin et al. (1984) make the more controversial claim that the reductionist and determinist thinking present in fields such as sociobiology and behavior genetics fuels conservative and anti-egalitarian political agendas.

Where do these purported reductionist tendencies in the field of behavior genetics come from? Some analysts imply that the most likely explanation is that behavior genetics practitioners themselves hold flawed ideas about the function of genes and their power to predict human behavior. For example, in a recent article critiquing several experiments on maternal nurturing behavior that used knockout mice, Rosoff (2010) argues that "by geneticizing phenotypes which may be immune to such causal attribution, [behavior geneticists] pervert their methodology in the vain pursuit of a concrete, discrete and objective answer to an inherently diffuse, massively multifactorial and subjective question" (p. 228). He points out that researchers rarely use causal language in their papers to describe the role of genes and that most practitioners acknowledge important roles for environmental influences, but he argues that "there remains a core commitment to the central role of specific genes" in the field (p. 202). Steven Rose (1997) calls reductionism the "ideology" of neurogenetics (p. 273). He points to the field's historical origins in eugenic science as one of the forces shaping the reductionist approaches of the field today. (S. Rose,

1997)[1]

To provide a corrective to what they view as flawed assumptions in the epistemic underpinnings of the field, some critics have adopted a "debunking" stance when commenting on behavior genetics research. Rosoff, for example, writes that "demonstrate the emptiness of any scientific or lay claims to strong deterministic claims between these genes and the behaviors they supposedly 'cause'" by " laying bare [the] history and complexity" of experiments he is examining (2010, p. 204). Freese and Powell (2003) argue that while such critiques of behavior genetics can serve to re-emphasize the limitations of the field's methods, many such commentaries are simply "tilting at twindmills [sic]," reiterating points that behavior geneticists already acknowledge as the limitations of their own work (p. 130). In his study of human genetic research on schizophrenia, Hedgecoe (2001) finds that the dominant narrative in the research community is not one of hard-line determinism, but what he terms a narrative of "enlightened geneticization" that acknowledges the importance of the environment while subtly privileging the importance of genes. Nikolas Rose (2007) argues more broadly that many sociologists misunderstand the style of thought present in twenty-first century genetic research. Contemporary biomedical research may operate in the "flattened" field of genes, molecules, and brain circuits, but he argues that "the truth discourse of contemporary genomics no longer sees genes as the hidden entities that determine us" (p. 15). For example, he argues that analysts who suggest that genetic research represents a return to eugenics do not acknowledge the differences between conceptions of gene action in the first half of the twentieth century and today. While early twentieth century eugenics research was grounded in the idea of control over certain immutable characteristics, biological research in the twenty-first century is based on ideas of optimization and personal control of health risks. Rose argues that this new form of biomedical thinking is no less deserving of critical scrutiny, however,

[1] See also Paul, 1991 on connections between the early history of behavior genetics and eugenic science, and Lewontin, 1992 on genes as "ideology" in the biological sciences.

since the supposed malleability of "nature" leaves it open to new forms of intervention in contemporary society.

Several recent studies of animal models in the neurosciences have also suggested that the narrative of contemporary behavior genetics research cannot be characterized as straightforwardly "reductionist," even though scholars still argue that critiques about the potential dangers of offering genetic explanations for human behavior are justified (Cwiartka, forthcoming; G. Davies, 2010; N. Rose & Rached, 2009). They note that the animal behavior genetics literature already identifies many of the potential theoretical and interpretational issues raised by the field's critics, and that behavior geneticists even demonstrate some of these issues through their own research. G. Davies (2010) and Cwiartka (forthcoming), for example, both cite the multi-sited study executed by Crabbe et al. (1999) as one instance where researchers themselves acknowledge the importance of environmental factors and the limitations of genetic studies of behavior. As I have demonstrated in previous chapters, the researchers in the Smith laboratory also assert that they believe that the classically reductionist formulation of a gene that codes for a behavior (or even *genes* that code for a behavior) is patently false, and that the genes associated with behavioral disorders are likely to behave in "complex" ways that depend on their relationship to other genes and environmental factors. The Smith laboratory researchers train new practitioners to recognize "complexity" at the genetic, phenotypic, and especially environmental levels; and the amount of time, attention and expense that researchers at Western devote to carefully controlling experimental settings and the way that they talk about the future of their field suggests that they are anything but certain of finding genes that "cause" alcoholism.

Analysts have also argued, however, that this shift in discourse about the nature of behavior is not necessarily associated with a concomitant shift in the methods that researchers use to investigate it or the data that they produce. Several scholars have argued

that reductionist approaches continue to dominate genetic research even in the midst of a growing consensus that genes may have limited value in predicting the presentation of complex diseases. In his study of the deCODE project in Iceland, Pálsson (2007) observes that hunting for genes associated with complex disorders is a "tiresome job" full of frustrations and disappointments, and yet the "paradigm of gene centrism" persists in spite of common awareness of these difficulties (p. 40). Keller (2008) likewise notes the "persistence of unproductive debates about the relative importance of nature and nurture" even in the midst of widespread agreement that the nature-nurture dichotomy is essentially meaningless (p. 117). Lerner (2006) describes reductionist and determinist ideas in behavior genetics as undead entities that have continued to "rise from the grave" no matter how many argumentative nails are driven into their coffins (p. 336–337).

Some analysts attribute this phenomenon to factors such as media reporting and a culture of genetic essentialism (Conrad, 1999b; Conrad & Weinberg, 1996; Nelkin & Lindee, 1995), or the influence of markets and start-up companies hoping to exploit value of genetic information (Pálsson, 2007; Keller, 2008). In her exploration of mouse research on aggression, Cwiartka (forthcoming) argues that the practical limitations of working with mouse models might provide one explanation for why animal behavior geneticists continue to offer reductionist answers to complex problems. Drawing on Fujimura's (1987, 1996) concept of "doable" problems in research and Pickering's (1995) argument about the "mangling" of researchers' intentions with materiality, Cwiartka conceptualizes both the mice as experimental objects and the narratives that scientists can produce about these objects as "constrained" by "available experimental resources and the inherent material capacities of the object" (p. 7). Cwiartka argues that acknowledging the limitations of animal models and the data that these experimental systems can produce may be the best way to avoid problematic applications of behavior genetics research to social policy.

Scholars have also pointed to the role of metaphors in supporting reductionist ideas

about genetics. In their seminal work on metaphor, Lakoff and Johnson (1981) argue that metaphors are more than just "rhetorical flourishes"; they posit that metaphorical comparisons are basic to human thought processes. Pointing to the pervasive use of metaphors in daily life, they argue that metaphorical thinking structures everyday activity by allowing us to understand and experience one kind of thing in terms of another. Lakoff and Johnson describe metaphors as having certain "entailments" that highlight some features of an object or experience and mask others, and that allow the metaphors to be extended in some directions more easily than others. Looking at the role of metaphors in the sciences, Kay (2000) explores in depth how the idea of DNA as a language and the genome as the "book of life" oriented genetic research programs in the 1950s and 60s, and has become so deeply rooted in scientific and cultural vocabularies that it is now easy to understand the genome as literally spelling out the essence of a person.[2] Pálsson (2007) similarly argues that the metaphor of genetic relatedness as a "tree" reinforces deterministic and essentialist classifications of people. In his study of metaphor in behavior genetics, Nordgren (2003) found that behavior geneticists also use metaphors pervasively in both textbooks and scientific publications, suggesting that metaphors are not only used for pedagogical purposes but play a more basic role in scientific thinking in the field. Nordgren finds that the authors of the texts he examined explicitly reject metaphors that imply deterministic gene action, such as the formulation of genes as "programs" or traits as "hard-wired," and favor metaphors that portray genes as agents that can "produce" traits and "interact" with environments.

I suggest that animal models and their epistemic scaffolds can also be thought of as metaphorical structures that allow researchers to speak about, think about, and produce data about the human in terms of the mouse. As I summarized in the preceding chapter, producing knowledge with model organisms has been described by several analysts as

[2] See also Keller, 2002 on information metaphors and understandings of gene action.

a process of using information generated using experimentally tractable organisms to explain the less understood features of the human. The epistemic scaffolding of animal behavior genetics models links the mouse and the human in ways that allow researchers to justify research programs that use animals to generate knowledge about specific aspects of human behaivor and biology. The configuration of this scaffolding also has certain "entailments" that enable researchers to easily emphasize aspects of human behavior or extend the mouse-human metaphor in particular directions. In the case of the elevated plus maze, for example, the pharmacological and ethological arguments each impart different visions of the mouse and of human anxiety. Particularly at the lower levels of the epistemic scaffolding, the pharmacological ladder masks many aspects of mouse behavior and focuses only on how it responds to drugs. Whether the mouse is experiencing something like anxiety is irrelevant; all that matters is that mice behave in a certain way (i. e., spending more time in the open arms of the maze) when they're given anxiolytics, and that this behavioral change predicts drugs that will be effective in humans. The mouse that is represented in the pharmacological ladder appears as little more than a lively biological detector for the effects of drugs. This cross-section of the epistemic scaffolding also presents a very flat picture of human anxiety, where anxiety is defined by response to anxiolytic drugs. The ethological ladder, in contrast, draws more information from the natural history of the mouse into the epistemic scaffolding. Although the mouse might not experience anxiety disorders as humans experience them, the ethological explanation assumes that the mouse is capable of experiencing some sort of fear or stress that is related to what humans identify as anxiety disorders. The situations that might cause stress for a mouse are not necessarily the ones that would cause stress for humans, but while the anxiety-provoking stimuli might differ, the innate responses of both mice and humans are assumed to be similar. This cross-section of the epistemic scaffolding supporting the elevated plus maze depicts human anxiety disorders as rooted in ancient evolutionary

instincts, shared with mice, that are triggered too often or in the wrong situations in day-to-day human life.

The remainder of this chapter examines the metaphorical entailments of two different research projects: new binge drinking models developed in the Alcohol Research Consortium, and a database to collect phenotype information on mice developed by the Mouse Phenome Project. In my examination of these cases, I do not presuppose that modeling the genetics of behavioral disorders using animals automatically leads to the reduction of human behaviors to genes, as the genetic reductionist critique might suggest. Instead, I look at how researchers use these models as resources to talk about human drinking, and draw in other kinds of culturally available information about human drinkers as they work on the epistemic scaffolding of the models. The models developed by the ARC are designed to allow researchers to understand binge drinking humans in terms of the genetics and biology of mice, but I explore how the developers of these models also use them to talk about the environmental contingencies and multiple motivations that might lead human drinkers to drink excessively. Similarly, the Mouse Phenome Project is designed to help researches investigate the relationships between different genotypes of inbred mice and their behavioral and physiological traits, but the protocol information collected in the database also offers resources for talking about the clinical importance of human environments.

3.2 Developing an Animal Model for Binge Drinking in the ARC

In the early 2000s, the National Institute on Alcohol Abuse and Alcoholism (NIAAA) funded several large consortium projects that I will refer to in this chapter as Alcohol Research Consortiums or ARCs. These consortium projects were a novel type of collaborative research arrangement within the NIAAA, designed to bring together alcohol

researchers with various disciplinary backgrounds to study aspects of the same model and to recruit new researchers to studying alcohol and addiction. Consortium researchers used a variety of animal models (including mice, rats, flies, and monkeys) to study different drinking behaviors (such as binge drinking, drinking behavior after withdrawal, drinking and stress, and chronic drinking). According to an early progress report, these projects represented "the largest ever concerted effort to collect and integrate scientific data on neuroadaptation to alcohol consumption," using an approach that "combined study of animal behavioral models with molecular, cellular and systems-level measures of brain function."

The alcohol research consortiums funded by the NIAAA were proposed as a novel solution to a problem of knowledge production in the behavior genetics field that I noted in chapter 1. Since researchers in the alcohol behavior genetics field saw animal experimental systems as only partially stabilized, the many kinds of variations in testing protocols and testing environments employed by different researchers made it difficult to compare studies and arrive at conclusions about the role of particular genes. Raymond, an official at NIAAA, recalls that when he was involved in evaluating the neuroscience portfolio at the NIAAA in the mid-1990s, he observed the problems caused by variation in testing protocols in alcohol research:

> One thing that became apparent, at least to me, was the fact that people were working with their little unit in terms of research. So if they had one model, people would work on that model. Another guy would work on the same model but change things a bit, both of them would modify something, and if you looked at the data, the results coming out, you could never correlate them, even say supposing you were looking at something as simple as a two bottle choice, okay? Some people do 24 hours, some people do limited access, you know different schedules, and you could never correlate one or the other. And you know, pieces of data were coming out and we were not confident that we could make a legit statement.

The alterations that many researchers made to different test paradigms made it hard to generate definitive statements about experimental outcomes. Even relatively straightforward

experiments, like the classic "two bottle choice" experiment (where researchers give mice a bottle of water and a bottle of alcohol solution and measure how much they drink) were subject to these variations: While one researcher might leave the alcohol and water bottles on the cages for a full 24 hour period, others might run the two bottle choice procedure by giving mice access to water and alcohol for only part of the day.

One solution to this problem, proposed by Raymond and others in the alcohol research community, was to gather together researchers from several sub-specialities into consortium groups that would each work on one model. By studying a single model using the techniques of different specialists— such as behaviorists, molecular biologists, electrophysiologists, and neuroanatomists—the NIAAA hoped to accelerate the accumulation of knowledge in a few key areas of alcohol research. In the initial years of the consortium group that I followed, researchers decided to focus some of their efforts on modeling "binge drinking" behavior, a particularly difficult problem in animal research. "Binge drinking" is a pattern of consumption that is widespread amongst human drinkers, at least in the United States. The NIAAA defines binge drinking in humans as "a pattern of drinking alcohol that brings blood alcohol concentration (BAC) to 0.08 gram percent or above. This typically happens when men consume 5 or more drinks, and when women consume 4 or more drinks, in about 2 hours" (National Institute on Alcohol Abuse and Alcoholism, 2004). Some researchers estimate that approximately 75 percent of the alcohol consumed by adults in the United States is consumed in "binges," and this figure may be as high as 90 percent for adults under the age of 21 (Office of Juvenile Justice and Delinquency Prevention, 2005).

Despite the social significance of binge drinking, few researchers had attempted to model this pattern of drinking using mice because of a property of mouse behavior and biology that is well known in the field—namely, that mice don't like to drink. When given a bottle of alcohol solution, only a few strains of mice will drink a good quantity, most

will drink only a little, and some will drink none at all. The researchers I interviewed offered various reasons why this might be the case: Mice are "neophobic" and are fearful of new things in general and of new foods in particular, since they have no vomiting reflex and can't purge something once they have ingested it. Alcohol tastes and smells bad, perhaps especially so to certain strains of mice, and mice might avoid it because of these "adversary oro-sensory properties." Mice are "prey animals," and might avoid getting intoxicated enough that it decreases their ability to detect and escape from predators. Some researchers suggested that mice might not even be capable of drinking to intoxication because they metabolize alcohol so quickly, and therefore would have to consume large quantities in order to actually get drunk.

Even though mice are easy to use in a laboratory setting and are excellent tools for genetic studies because of their well-characterized genomes, the general tendency of most mouse strains to avoid alcohol makes them difficult subjects for alcohol genetics studies. Alcohol researchers have a variety of existing techniques for overcoming this difficulty and getting more alcohol into the bodies of their research subjects: Sweetening and flavoring alcohol solutions, offering food along with alcohol, or providing alcohol bottles at regularly scheduled times of the day are a few of the techniques that will increase the amount that mice will drink. In cases where these procedures are too time-consuming or still don't produce high enough rates of alcohol consumption, researchers may use methods like housing mice in specially designed chambers and letting them breathe alcohol vapor, or injecting alcohol directly into their abdomens. None of these existing methods, however, seemed especially promising to members of the consortium as models for binge drinking. Protocols that increased the amount mice drank still generally didn't get mice to drink large amounts of alcohol in short enough periods of time, and techniques such as injecting alcohol into mice were a "tough sell" as a good model for the human condition, since, as one researcher put it to me, "nobody sits around doing IV alcohol."

To address this gap in the field, the ARC researchers proposed the development of new protocols that would induce mice to drink enough alcohol that they would actually become intoxicated: the "timed drinking" model and the "evening drinking" model. The timed drinking protocol used a combination of limiting access to water and offering alcohol at regularly scheduled times to get mice to drink large quantities of alcohol. In this procedure, researchers give mice a water bottle for ten hours a day instead of having water available at all times. During the other 14 hours when the mice have no water bottle on the cage, researchers give the mice a bottle of alcohol for half an hour. Researchers in the ARC showed that limiting the amount of time that mice had access to water and offering a bottle of alcohol for a short period was enough to get mice to drink substantial amounts of alcohol in the half an hour window. The second protocol, the evening drinking protocol, used patterns in the eating and drinking habits of mice to encourage heavy drinking. Mice tend to consume most of their food and water a few hours after they wake up, and researchers showed that if they gave mice a bottle of alcohol early in the "dark cycle" (the time of day when mice are active), they will typically drink large quantities. Mice in the evening drinking protocol were allowed to drink as much water as they wanted for the entire day, except for a two hour period early in the dark cycle when their water bottles were switched for bottles of alcohol solution.

Consortium researchers conducted initial experiments to see how much mice would drink under these new experimental conditions and how reliable the results were, and the consortium as a whole debated the relative merits of each model and which one offered the best prospects for studying human binge drinking. One of the main goals of the ARC researchers was to develop models that would be easy to use in a laboratory setting and compatible with existing laboratory practices; or in other words, to generate extremely stable protocols that could be used as "technical objects" by other researchers and integrated into existing experimental systems.(Rheinberger, 1997) The models were designed to be

resources that could be extended to other researchers outside of the consortium to use in setting up future experiments on heavy drinking. To ensure that the models were as stable as possible, consortium researchers conducted a series of experiments to assess how various factors impacted the mouse's drinking, such as the time of day that alcohol bottles should be introduced and how long they should be left on the cage to induce the greatest amount of drinking. Linda, a consortium researcher, recalls that this aspect of developing the new models was quite successful and that she found them to be very reliable tools when she first started using them in her laboratory:

> That's the one thing that I have to say is that the beauty of these models, both of them, [timed drinking] and [evening drinking], is you can take that model and apply it in any place, and they will at least drink 0.08 %. And we're getting numbers that are bang on, we're getting bang on with what [other researchers in the ARC] have found. And we've got undergrads running the experiments, we've got grad students, we've got people who are probably hung over when they're doing the experiments, and they're getting the same type of data. So it is a beautifully robust model.

One of the key features that made the models convenient to work with in a laboratory setting was the fact that researchers could anticipate when mice would drink in both of the models and schedule their experiments accordingly. Charles recalls that this was the feature that made him excited about working with the evening drinking model:

> There have been different drinking models going back forever. Well, the cool thing about this was you knew precisely when the mouse is going to drink. And therefore you knew precisely when to administer some treatment to manipulate the drinking, as opposed to a two bottle choice kind of situation where the animal is drinking, you know, around the clock. You can go in and give a microinjection any time, but is it the appropriate time, and would you really expect to see any effect on the drinking that you could measure?

Instead of leaving a bottle of alcohol on the mouse cage all day and hoping to catch a mouse when it had been drinking, researchers could schedule experiments directly after the drinking window for maximum physiological effect. Using these protocols, researchers could reliably produce large groups of mice who had just consumed large amounts of alcohol for use in future tests.

Another goal of the consortium researchers was to develop protocols that would induce mice to drink enough that they would achieve high blood alcohol levels in the drinking windows. The selection of blood alcohol levels as the standard for assessing the success of the new models was theoretically important for the alcohol research field, and in interviews ARC researchers told me that one of their main contributions was to demonstrate to the research community that mice and rats could reach intoxicating blood alcohol levels in drinking experiments. Prior to the development of the timed drinking and evening drinking models, researchers primarily measured the amount of alcohol that mice drank over a given period of time, and most researchers assumed that rodents were either drinking too slowly or metabolizing alcohol too quickly to actually get drunk. Larry recalls that some members of the consortium raised this objection in initial discussions about whether it was even possible to create a binge drinking model. He says:

> [Some members] had trouble getting their heads around the fact that mice drink past metabolism, because the old story put out there was that mice only drank for calories. You see what I mean? They never drank to get intoxicated, and what [we] showed is that they do drink to get intoxicated. Because they're intoxicated!

Testing the blood alcohol levels of mice directly after the drinking windows in the timed drinking and evening drinking models offered a way to demonstrate to the research community (and even to some members of the consortium) that it was possible to get mice to drink to intoxication, and therefore to model human intoxication using mice. Through a series of experiments with the new models, consortium researchers showed that it was even possible to produce mice who drank to an intoxicating blood alcohol level of 0.08 percent in protocols that depended only on drinking alone, rather than more invasive procedures like giving mice injections of alcohol or letting them breathe alcohol vapor.

3.2.1 Reductionism in Modeling Binge Drinking

All of these experiments, arguments, protocols, and publications can be thought of as part of the process of building the epistemic scaffolding to support research into human "binge drinking" using these models. Demonstrating that mice in these models would drink enough to reach a blood alcohol level that corresponded with the NIAAA's definition of binge drinking in humans provided the foundation for a new epistemic scaffolding to support research with mouse models for "intoxication" rather than just "alcohol preference." In addition to these theoretical objectives, the construction of these models was also shaped by practical considerations, such as how easy these protocols would be to run in a laboratory setting and to use with existing experimental systems. Once established, researchers intended that these protocols would act as "platforms" for future research projects both within the consortium and for others in the alcohol research community (Keating & Cambrosio, 2003). Researchers could use these established models to look for genes that predisposed the mice to drink heavily, neuroadaptations that happened as a result of heavy drinking, or drugs that modified a mouse's tendency to drink.

From the very beginning, the ARC researchers who are involved in formulating these new models of binge drinking are also creating new sets of metaphorical relationships that link the mouse and the human. In choosing a 0.08 percent blood alcohol level as the criterion for a successful model, researchers are selectively using information about human drinkers to highlight features of the ideal drinking mouse. This enables researchers to portray some models (such as existing mouse models, or mice in general) as "bad models," and to argue that their models are an advance over existing models because they capture the key feature of high blood alcohol levels present in human drinkers. When Frank suggests that existing models for studying intoxication are bad models because "no one sits around doing IV alcohol," he is also arguing metaphorically. He makes his case for the shortcomings of these models by extending the metaphor of experimental mouse

as human in a direction that is humorous because it does not fit with commonsense ideas about human drinking.

What kinds of entailments do the emerging epistemic scaffolds of the timed drinking and evening drinking models have for how researchers talk about humans? Do the analogies between the mouse and the human embedded in these scaffolds make it easier for researchers to talk about human drinking in genetic terms? Researchers say that their eventual aim in developing these new models is to use them to produce genetic or biological data about human binge drinking, and in this sense the approach of the ARC researchers could be fairly described as a "reductionist" approach that seeks to explain a human behavior in terms of biological properties, especially genes. Blood alcohol levels, the link at the base of these epistemic scaffolds, offers a strong foundation for building new research programs because it is a quantifiable, biological marker that can be used across species. But this narrow criterion also excludes many arguably important features of human binge drinking, such as how behavior is altered by drinking large amounts of alcohol or the social settings that contribute to heavy consumption. In his discussion of "drunk driving," Gusfield (1981) points out that blood alcohol levels have also been used as the basis for public policy because it offers a way to cleanly differentiate those who are "drunk drivers" from those who are not, despite the fact that blood alcohol levels are not isometric with intoxication.

This section explores how researchers themselves envision their research programs with animal models and how they will contribute to understanding human drinking behaviors. I argue that many of the arguments made by critics of the "disease concept" of alcoholism are reflected in the alcohol research field's own conceptualizations of alcohol as a "complex" disorder. While researchers, unlike some critics, do assert that alcoholism has a biological basis, they also argue that there are cultural aspects of drinking behavior that cannot be captured in animal models. Animal behavior geneticists describe their approach to

investigating alcoholism as an "intentionally reductionist" approach that brackets out many features of human alcoholism in order to produce information about biological factors, and as an approach that is complimentary to other ways of understanding alcoholism.

Numerous commentators have argued that understanding alcoholism through biology is neither the only nor the most productive way to understand these behaviors. The idea that alcoholism is a biological, genetic, or a "brain-based" illness is culturally and historically specific (Levine, 1978/1985; J. W. Schneider, 1978; Conrad & Schneider, 1980; Room, 1983; Nelkin & Lindee, 1995; Campbell, 2007, 2010). Conrad and Schneider (1980) have described the shift towards treating heavy drinking as an illness as one example of a more general trend towards the "medicalization" of social problems in contemporary Western society. They argue that cultural understandings of habitual drunkenness in the United States have shifted: What was once seen as a "sin" or a moral failing to be addressed by the church or the legal system is now largely viewed as an "illness" to be treated by self-help groups and medical professionals. They point to the involvement of powerful actors and interest groups that supported the idea of alcoholism as a disease, such as the self-help organization Alcoholics Anonymous (AA). AA's publications portray alcoholics as people who are sick, an idea that Conrad and Schneider argue was "appealing both to drinkers and to those charged with their care" (p. 89). Ethnographers have also described some of the interactional processes that sustain cultural ideas about addiction and alcoholism (Becker, 1953; MacAndrew & Edgerton, 1969; Reinarman, 2005; Gusfield, 1986). The process of becoming an alcoholic or an addict, they argue, requires more than just the ingestion of drugs; it requires a process of learning to understand their experiences as drug-related, learning to behave as someone who is under the influence of drugs, and learning to identify themselves as "addicts." Reinarman (2005) argues that the fact that the idea of addiction as disease resonates with many addicts should not be surprising, since drug users' experience of drug taking are articulated through culturally available ideas

about addiction.

The movement to treat heavy drinking as an illness rather than a social problem has been called the "disease concept" of alcoholism, after alcohol researcher E. M. Jellinek's (1960) text entitled *The Disease Concept of Alcoholism*. Room (1983) describes the disease concept, especially as it was articulated in the 1950s and 1960s, as a set of several interrelated assumptions that include: 1) the premise that there is a well-defined singular entity called "alcoholism," which some people have and others don't; 2) that those who have "alcoholism" will always be different with respect to their drinking patterns and should never drink again; and 3) that "alcoholism" should be thought of as a disease in and of itself, and not as a symptom of another disease (Room, 1983, p. 54). Alcoholics Anonymous' famous slogan "one drink, one drunk" epitomizes this way of thinking about alcoholism as a disease with a defined entry point and a defined pathway: When susceptible individuals encounter alcohol, the disease is flipped on like a switch and alcoholics quickly find themselves unable to control their drinking (Fingarette, 1988, p. 20). The idea that alcoholism is a disease akin to other kinds of psychiatric illnesses is today widely accepted in scientific, medical, and policy communities. A widely cited definition of alcoholism developed by a committee of the National Council on Alcoholism and Drug Dependence and the American Society of Addiction Medicine defines alcoholism as "a primary, chronic disease with genetic, psychosocial, and environmental factors influencing its development and manifestations" that is often "progressive and fatal" (Morse & Flavin, 1992, p. 1012).

Critics of the "disease concept" have argued that this formulation is too simplistic, and that these behaviors could also be explained in sociological terms rather than biomedical ones. In his book *Heavy Drinking: The Myth of Alcoholism as a Disease*, philosopher Herb Fingarette (1988) argues that "alcoholics" are not a unitary group of disease sufferers, but are people who turn to heavy drinking for many different social and cultural reasons. Contrary to what the disease concept of alcoholism would predict, he points out that

drinkers have a variety of drinking patterns, and even some long-term heavy drinkers can control their drinking under certain circumstances. He argues that the diversity of drinkers and patterns of use discredits the idea of alcoholism as a disease:

> There are, in short, many kinds of heavy drinking that arise from many different causes and produce many different patterns of associated problems … Instead of looking at heavy drinkers as victims of some wayward gene or a physical abnormality, we can now see them in a truer light: as a diverse group of people who for diverse reasons are caught up in a particularly destructive way of life. Although this depiction is messier than any single-factor theory, it has the advantage of being true to the observations of clinicians, and to those of many heavy drinkers, and their families and friends (p. 65–66).

He also cites studies that show that only a percentage of children of alcoholics actually become alcoholics themselves as evidence that alcoholism is not a biologically-based disease. Fingarette's invocation of an "alcohol gene" links the idea of alcoholism as a unitary entity to expectations for a single biological cause. Without a well-defined group of "alcoholics," he argues, it makes no sense to assume that there is an underlying biological cause. Steven Rose (1995) makes a similar critique of the behavior genetics approach. He suggests that "complexity is hard to deal with within the neurogenetic agenda," but that "if there is one single thing called alcoholism, for instance, then it becomes appropriate to seek a single causative agent" (p. 381; see also S. Rose, 1997). He argues that one of the flaws of behavior genetics in general is that researchers proceed by inappropriately lumping together different specific instances into global categories for the purposes of investigation.

These ideas about the heterogeneity of alcoholism have been so thoroughly absorbed into the mainstream of contemporary alcohol research, however, that they might appear to alcohol researchers to be at best an accurate assessment of the nature of complex disorders or at worst an argument against a straw man definition of alcoholism as a disease. As Nikolas Rose (2007) observes, for many biomedical research communities the idea that something is a "disease" no longer requires that there is a unitary biological cause that

underlies it, or that whatever biological causes exist will determine the pathway of disease in any strict sense. Raymond, for example, also describes alcoholism as a heterogeneous disorder with multiple pathways that might lead to alcohol dependence. He says:

> Usually in terms of alcoholics, it takes a good 15, 20 years to become an alcoholic. But then you get a different set of alcoholics, those who start very young and become dependent by the time that they're 25. That's a different cohort of people. Clinically, if you look at the disorder it's very heterogeneous. The trajectory to alcohol dependence is varied, there are about four or five different ways, and no two alcoholics are the same. While in cancer, the tumor will be the same, right? You have a breast tumor, it's a tumor.

Raymond's description of alcoholism as a heterogeneous disorder with multiple pathways to dependence in many ways resembles Fingarette's (1988) critique of the disease concept of alcoholism. Raymond describes alcoholism as more varied in its presentation than even other kinds of complex diseases such as cancer. Alcohol researchers assume that the trajectory to alcohol dependence includes biological factors, but they also acknowledge a role for psychological and sociocultural factors. Windle (2010), for example, describes his approach to studying binge drinking in human adolescents as a "multilevel developmental contextual approach" that uses measures of values and beliefs, interpersonal functioning, and antisocial behaviors, as well as biological indicators to predict who will develop patterns of heavy drinking later in life.

If there are many ways in which a human can become an alcoholic, then it follows that there might also be many ways in which to model that process using animals. Raymond describes the idea that there is (or should be) a single animal model for alcoholism as a common misperception that those outside the field hold about what modeling means in the context of alcohol research. He explains:

> There's no one animal model for alcoholism. If you did, you would have drunk rats, but you don't. They don't like the taste of alcohol, you have to initiate them and get them to drink, right? So we're clear on that. It took people from outside the field a lot to understand that, that the animal models model one facet of the continuum that you see in alcohol dependence.

A single animal model, in Raymond's description, cannot capture all of the facets of human drinking, but it could still be useful for investigating more narrowly defined features of alcoholism. His comments suggest he thinks there are practical reasons why an animal model for "alcoholism" is not possible, such as the reluctance of rodents to drink, as well as theoretical reasons, such as the diversity of drinkers who fall on the "continuum" of alcohol dependence.

Many alcohol researchers, including Dennis, do not believe that researchers will ever be able to capture all of the important features of alcoholism even with multiple animal models. Dennis points out that the DSM-IV outlines a list of seven criteria for diagnosing alcohol dependence, ranging from withdrawal symptoms to giving up important social and occupational activities because of alcohol use. He considers a few of these characteristics to be well-modeled in the animal research community, such as tolerance to alcohol and withdrawal. Modeling other features of the disorder, such as craving for alcohol, is something that he believes animal researchers have yet to do successfully. On the whole, Dennis thinks that the majority of the criteria for alcohol dependence as outlined in the DSM are "pretty tough" for animal models to capture. He says:

> Five of those seven symptoms are things like losing your job or persistently continuing to drink even in the face of evidence that your health is falling apart. You know, the doctor telling you you're killing yourself, your liver's getting trashed and you keep drinking … your relationships fall apart, you wind up in jail, you're obsessed with getting the drug. Those are all very human symptoms that are at the core of the disorder.

DSM criteria like "[continuing] alcohol use despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the alcohol (e. g., continued drinking despite recognition that an ulcer was made worse by alcohol consumption)"are outside the scope of what Dennis thinks that researchers can reasonably expect to emulate with animal models.

In Dennis's opinion, the best way for animal researchers to contribute to understanding human drinking is by adopting an "intentionally reductionist" stance that breaks down the

complex disorder into discrete, biological characteristics that can be studied in animals. He explains:

> So if you try to model these complex things, most of us start by being reductionists, whether we want to or not. We start as obligate, intentional reductionists. I'll say, okay, one of the symptoms of alcoholism in humans is tolerance, so if you started out at 20 years old being able to drink 3 beers before you fell over and passed out, by the time you've been doing that for 7 or 8 years, your dose has escalated like crazy and now you can drink 6 beers before you fall over and pass out and that's one definition of tolerance. You can model that quite easily in animals in lots of ways, and so there's been a lot of research on trying to understand what changes in your brain as you become tolerant.

Rather than abandoning research into genetic contributors to alcoholism entirely because of the heterogeneity of human drinkers, many alcohol researchers advocate using animal models to study specific aspects of the disease spectrum of alcoholism. Instead of modeling the global disorder "alcoholism," researchers argue that animal models can represent different pathways to dependence, or stages in the disease, and would allow researchers to investigate specific questions such as how gene expression changes in the brain as drinkers develop tolerance to alcohol. Dennis's comments also suggest that the number of drinks a human drinker can consume before passing out is only one possible definition of tolerance, and other researchers might also contribute to the understanding of "tolerance" using models that begin from different definitions or measures.

To summarize, alcohol researchers themselves describe their approach to animal modeling as "reductionist," and assert that alcoholism is a disease rather than a deviant behavior or a coping style. In modeling binge drinking behavior, researchers in the ARC selected high blood alcohol levels as the feature that they aimed to reflect in their models, a biological criterion that provides a narrow view of human binge drinking. But, the consortium researchers describe their reductionist approach to modeling binge drinking as a pragmatic methodological solution to addressing a complex human problem. They argue that animal models do not capture the full spectrum of human drinking patterns or all of

the causes and consequences of alcoholism. The ARC researchers' introduction of the new models for binge drinking in the field shows that even some basic features of human alcoholism, such as the tendency to drink to intoxication, were not considered to be well modeled or even possible to model using rodents by the alcohol research community. Animal behavior geneticsts' portrayal of their models as partial, overlapping, and incomplete suggests that they believe that not all facets of human drinking behavior can or should be reduced to genes.

3.2.2 Reducing Humans to Environments

Genetic animal models for alcoholism are designed to allow researchers to talk about human drinking behaviors in biological or genetic terms, but in this section I explore some of the ways in which researchers in the ARC also use these models to emphasize the importance of factors other than genes. In particular, I examine how researchers talk about what motivates humans and mice to drink alcohol, and I argue that "motivation" is an area of the epistemic scaffolding of the timed drinking and evening drinking models that has flexible metaphorical entailments. Researchers characterize the forces that motivate mice in these models to drink as both biological and environmental, and they use these models as metaphors to emphasize how human drinking is structured by a person's surroundings as well as genes.

In addition to the shared similarities in blood alcohol levels between the mouse and the human, another feature supporting the use of these new protocols as models for human binge drinking was the "natural" and "unmanipulated" circumstances of the drinking situation. Compared to other models that used more invasive techniques like injections to achieve high blood alcohol levels, mice drinking in the timed and evening protocols appeared to be drinking with relatively little coercion. A mouse drinking to intoxication a few hours into the dark cycle offered a more convincing analogy to a human binge drinker

than other available rodent models. As Dennis put it, the models offered an advance over existing ones because in these models mice really did " drink until they fall over, just like college students do."

In many instances, the researchers involved in developing these models described the forces that are motivating mice to drink in the timed drinking and evening drinking protocols as biological properties. When describing the reasons why mice in the evening procedure drink large amounts, Linda attributes this pattern of drinking to the deliberate choice on the part of researchers to introduce alcohol to mice at times of the day when they have an "internal inclination" to drink. She explains:

> Three hours into the dark cycle, that was not selected willy nilly. There's studies on feeding and drinking behavior and the circadian rhythm, and that is the time when typically mice and rats consume most of their food and water. So that's like an innate cycle that they have for regulating their intake, so when you present booze at that time point, the animals already have an internal inclination to eat and drink at that time. So if that booze is the only thing that's available…that's what's going on with these models, right? One bottle, boom. They're going to drink, because that's what their internal clock is telling them to do.

Linda attributes the success of the procedure to researchers' ability to align testing schedules with the mouse's "innate cycles" of food and water intake. Placing the factors that motivate mice to drink in the evening drinking protocol on the side of the "internal" and the "innate" portrays these factors as part of the biology of the mouse rather than the mouse's environment, since these types of descriptions tend to be associated with nature rather than nurture.

Another way that researchers represent the mouse's drinking in these models as a biological inclination is by talking about the experimental situation as "natural" or "unmanipulated." Charles, a researcher who joined the consortium several years after its inception, recalled that he was especially excited to work with the evening procedure because it seemed to him to be a more valid model that employed fewer "manipulations" than other available models. He says:

> I would say in a lot of models, you know the intragastric intubation approach is going to be more stressful and therefore you have a stress, a potentially new stress confound. What I like about drinking in the dark is that there is no manipulation to the mouse whatsoever. Now that's in a mouse that I haven't cannulated,[3] but there is no manipulation to the mouse whatsoever.

He contrasts the evening drinking model with another model for generating heavy drinking mice, intragastric intubation. In this model, researchers train mice to press a lever that pumps alcohol directly into their stomachs so that the mice can "drink" alcohol without ever having to taste it. But Charles argues that surgeries and technical equipment required to make this model work introduce new problems and stressors that may confound the experimental results. He suggests that unlike other models, mice in the evening drinking model just seem to drink naturally with few interventions on the part of the researcher. His portrayal of the evening drinking experimental situation as free from external influences (at least until he cannulates the mouse) suggests that the mouse is drinking because it is intrinsically motivated to do so.

In some instances, researchers describe the motivations driving the behavior of the mouse as both biological and environmental. In a paper outlining the timed drinking procedure, the authors attributed the success of the protocol to the combined effects of the "natural" preferences of the strain of mouse that they used for drinking alcohol and the effect of "conditioning" those mice to drink on a schedule. By initially restricting the amount of time during which mice have access to water and offering them an alcohol bottle for a short period every day, researchers can turn mice who have a slight preference for alcohol into mice that will drink intoxicating amounts. This description portrays the drinking mouse in the experimental setting of the timed drinking protocol as controlled by both biological and "learned" factors. The researchers who developed the timed drinking model suggested that the mouse's urge to drink might come from a genetic predisposition

[3] "Cannulating" a mouse is the process of inserting a small tube into a specific brain region of the mouse so that drugs can be delivered to that region or researchers can sample neurotransmitter levels from that region of the brain.

to drink alcohol as well as from external features of the experimental situation.

This brief description of the way that researchers talk about the mouse in the timed drinking and evening drinking protocols suggests that these models have mainly biological entailments that would encourage researchers to conceptualize the human drinker in biological terms, but researchers also referred back to humans through the lens of these models in more flexible ways. In my interviews and conversations with ARC members about the validity of these models, researchers frequently made parallels between human drinking and mouse drinking in the timed and evening protocols that suggested environmental rather than genetic causes for drinking behavior. Larry, for example, explained rat drinking in the evening model as shaped by certain contexts and conditions that allow drinking to progress further than it normally might, and compared these to similar situations that human drinkers might face. He says:

> I mean, the intoxication part, as with humans, requires certain contingencies to engage in. Sometimes, you know, I don't know why a rat in [the evening model] drinks so much, but they do. And presumably it's similar to one of us going to have a glass of wine before dinner but suddenly ending up with three, okay? So you're on an empty stomach, you may forget how many you've already had, or you like the feeling that you've gotten, or you've lost your appetite for the dinner because you're drinking, so you just keep drinking. I don't know why mice do it too, but they do.

Larry describes both human and mouse drinking as requiring certain "contingencies" for it to progress to a level where the drinker will become intoxicated. Using the rat's drinking in the evening model as a metaphor for human drinking, he suggests that a human drinking on an empty stomach might end up drinking intoxicating amounts because of the circumstances in which he or she is presented with alcohol. Both the human and the mouse, in Larry's description, seem compelled to drink by the structural features of their respective situations. In his analogy, the human drinker doesn't choose to have another glass of wine before dinner, he or she "suddenly ends up" having three. The human drinker, like the mouse, almost appears to be tricked into drinking more than he

or she intended, but not by biology. Here, the human drinker seems to be reduced to the features of his or her environment.

Linda makes an analogy that has similar characteristics in her speculations about why the timed drinking model works. When she teaches her undergraduate class on addiction biology, she says that she refers to the evening drinking protocol as the "happy hour" model and the timed drinking protocol as the "late for happy hour" model to explain to her students how they work. Talking about the timed drinking model, she says:

> I mean, who knows how much [the mice] are learning that the alcohol is only available for half an hour? Like, I call it the 'late for happy hour' model. You know, when you're late for happy hour, you order a whole bunch of drinks and you pound them all down, even though once you pay for them, they're not going to take them away. But in this case I don't know how much the animals are learning in terms of that half hour period.

Linda wonders whether the mice in the timed drinking model might be learning that alcohol is available for a limited time after several rounds of being presented with an alcohol bottle for only half an hour each day, and she compares this scenario to human drinking during happy hour at a bar. As in Larry's description of drinking on an empty stomach, the human drinker in this situation seems to be controlled by the structural features of his or her environment rather than by biology or genes. In Linda's description, the human drinkers are inclined to order and drink more than they normally would not because of genetic predispositions, but because drinks will only be available at special prices for a limited period of time. The way that she has extended the experimental metaphor to a human scenario makes the drinking human seem even more controlled by his environment than the drinking mouse: While at least a mouse in a timed drinking experiment would be correct in surmising that it has to drink quickly or the alcohol bottle will disappear, the human presumably knows that his drinks won't be taken away once he's paid for them—but drinks them in rapid succession anyway.

These examples demonstrate that the area of "motivations" in the epistemic scaffolding of both the timed and the evening models provides flexible resources for talking about

human drinking. While these models are designed to help researchers highlight biological and especially genetic factors that motivate or predispose humans to drink, researchers do not always use these models to describe a human that is controlled by genetic impulses. In teaching, informal conversations, and interviews, researchers also use these experimental situations as resources to make analogies to the human that highlight environmental and structural aspects affecting human drinking.

In using animal models to understand human drinking, researchers also transform images of alcoholism by highlighting or excluding features of the human drinker. One of the features of alcoholism that remains notably absent in the ARC researchers' discussions about binge drinking is the question of self-control or will. Commentators on alcoholism have pointed out that the "will" is central to cultural narratives about what it means to be an alcoholic rather than a normal drinker. Valverde (1998) argues that the first attempts to medicalize alcohol consumption did so by characterizing it as a "disease of the will," where consuming strong liquors or excessive amounts was seen as damaging to the drinker's willpower. She explores how the question of whether heavy drinking is a matter of a failure of the will or an uncontrollable condition of the body is also at the heart of contemporary debates about the status of alcoholism as a disease. The "loss of control" concept, for example, is a key diagnostic criterion for demarcating the boundaries between normal and pathological drinking in humans. Bailey (2005) contends that the very idea of addiction "rest[s] in a dualist conception of the mind–body relationship, it is conceptualized as a failure of the self in its imperative to exercise control over bodily desires and functions" (p. 539).

This formulation of addiction as a subjective state experienced as a loss of ability to control one's intake of alcohol is nowhere to be found in the consortium researchers' discussions of modeling binge drinking behavior. This omission is perhaps not surprising: as I argued in the previous chapter, researchers in the Smith laboratory and other users of

the elevated plus maze deliberately exclude the subjective experience or internal mental states of the mouse from the epistemic scaffolding of the tests. Speculations about the mouse mind are marked out as too "anthropomorphic" to be considered convincing, and researchers attempt to use animals as behavioral and biological models for human behavioral disorders without referencing these internal states. Crist (1999) argues that the deliberate erasure of the animal mind has a long history in behaviorist research traditions, where animals are treated as intentionless actors responding to environmental cues. In her study of experimental systems for investigating alcoholism, Campbell (2007) also argues that features of alcoholism like "craving," "control" or "desire" that fall somewhere between behaviors and mental states have often been avoided by those working with animal models. She argues that addiction researchers working with animal models in the 1920s and 30s chose to bracket off questions about the subjective experience of desire for drugs and focus instead on physiological effects like toxicity and withdrawal symptoms; but she also shows how the question of how to demonstrate "desire" for drugs in animal models has continued to resurface throughout the history of addiction research.

To summarize, researchers' comparisons between human drinking and rodent drinking in the timed and evening drinking protocols suggests that aspects of the epistemic scaffolding of these models have flexible entailments that allow researchers to think and speak about human drinking in both environmental and genetic terms. While the ultimate goal of developing these models is to use them to understand the genetics of binge drinking, the process of setting up these models also provides resources for thinking about environmental contingencies that might also structure human drinking. Understanding human drinking in terms of mouse models also reformulates representations of the human drinker. By separating out questions of will and self-control from the work of mouse modeling, researchers in the ARC structure portray drinking as a matter of conditioning and biology rather than a struggle between biological urges and self-control. Both pathological

drinking behaviors and potentially normal drinking behaviors (such as having a glass of wine at the end of the work day) are presented by researchers as controlled to some degree by the drinker's biology and environment.

3.2.3 Modifying Epistemic Scaffolds and Representations of the Human

Ongoing modifications and contestations about the configuration of the epistemic scaffolding can also have consequences for how the human is represented through animal models. This section examines the discussion around the eventual failure of the timed drinking model, and explores how representations of drinking human change as researchers argued about the validity of the model. Not all members of the ARC or the broader research community were satisfied that mice in the timed drinking model offered an acceptable analogy to drinking in humans because the protocol used water restriction to get mice to drink in large amounts. In response to this criticism, some ARC researchers questioned the implicit image of the human drinker as one who drinks only for intoxicating effects, and suggested that humans (like mice) might drink for taste, thirst, or out of habit as well as for the sensation of being drunk. These arguments highlight some of the ways in which ongoing work on the epistemic scaffolds of animal models can also modify represention of the human.

Both protocols developed by the ARC did induce mice to drink enough to reach the target blood alcohol levels, but not all of the consortium researchers were convinced that this meant they were good models for studying human binge drinking. Practical problems emerged when other members of the consortium began to take up the new models. The evening drinking model didn't seem to work so well for rats, or some strains of mice. Researchers accidentally and intentionally introduced variations into the protocols by using different lengths of drinking sessions or offering alcohol at different times of the day, making the experimental results coming out of the consortium laboratories once again

difficult to compare.

Some consortium researchers also raised conceptual issues about the validity of these models, focusing mainly on questions about motivation. Brian, for example, objected to the fact that both of the procedures that the consortium developed for modeling binge drinking were procedures where the animal had to drink the alcohol. In his opinion, drinking models were problematic because mice could be drinking (or not drinking) alcohol for many reasons. He explains:

> *Brian:* I have concerns that a lot of the drinking models aren't necessarily studying drinking that's motivated by intoxicating pharmacological effects.
>
> *NN:* What do you think it's being motivated by?
>
> *Brian:* Taste? Calories? I think [C57 B6 mice] drinking because it tastes sweet to them.

The strain of mice that were used to develop the protocols, C57 B6 mice, are widely used in alcohol research because they already rank as above average drinkers in many kinds of drinking tests. But Brian is concerned that they may simply be drinking because they have a different perception of the taste of alcohol or because alcohol is more nutritious than water, and not because they want to experience the sensations of being drunk. Likewise, he suggests that other strains of mice might be avoiding the alcohol bottle only because it tastes bitter, and not because they dislike being intoxicated.

Frank, another consortium researcher, thought that the timed drinking model in particular was a "disaster" because it relied on restricting the mouse's access to water to induce heavy drinking. When the model was first introduced in the ARC, he voiced concerns that the fluid deprivation procedures used in the protocol would produce too many behavioral and biological side effects in the mice. He explains:

> Rodents will do all kinds of stupid things when they're thirsty. They will lick at streams of air directed at their tongues for hours on end, and they'll work really hard to get access to those streams of air. When their fluid balance is upset, there's all kinds of side effects, and there's all kinds of genes that would be pulled in by an animal seeking to return to homeostasis.

Creating motivation through fluid deprivation, in Frank's opinion, was likely to result not only in erratic behaviors that should not be mistaken with a true motivation to drink alcohol, but also in biological changes that could interfere with future research efforts. When researchers eventually examined the genes expressed in mice who had been through the timed drinking procedure, for example, they might find genes that were activated because of the fluid deprivation and not because of drinking large quantities of alcohol.

When consortium researchers put forward funding proposals to the NIAAA based on the new models, they found that grant reviewers also shared some of these concerns. Researchers who tried to publish papers or propose research projects using the timed drinking model in particular encountered some skepticism from other alcohol researchers. After several grants that were based on the timed drinking model failed in the NIAAA review panels, the ARC consortium eventually decided to drop the model and focus only on the evening drinking model. Larry explains:

> Well, the reviewers didn't like it, that's the bottom line, and I personally had my own reservations. Some people liked it a lot because it engendered high drinking, but the problem is the animals are water deprived, and that's another motivational component that enters into it. And it's been a controversial issue in the alcohol field for a long time.

Members of the review panel argued that the mice in the model were probably drinking at least partially (and maybe entirely) because they were thirsty, and not because they wanted to ingest drugs. Larry suspected that the protocol also reminded the reviewers too much of an old model called "schedule induced polydypsia" that fell out of favor in the field because of similar concerns about motivation. He explains what the procedure looks like:

> If you start giving food to a food deprived animal, so a guy that's hungry, and you deliver it on a regular time point, like every sixty seconds, and you give them a sipper tube, the animal will start drinking huge quantities of water or sucrose or saccharin or anything, okay? If you put alcohol on, they'll drink that, and they'll drink large amounts, and become intoxicated. But that drinking is not specific to alcohol, it's specific to the motivations created by

> those contingencies … So it fell out of favor in the alcohol field for the wrong reason, and so in defense of it, it fell out of favor because when they took the contingency away, in other words they let the rats eat as much as they wanted and they didn't deliver the food pellets and then they gave them the alcohol back, they didn't keep drinking high quantities, or at least under the conditions that were set up in those days. And so people in that day said, "well, it's not a good model of alcoholism."

When researchers withheld food from a mouse, they found that the food-deprived mouse was apt to eat or drink large amounts of many different kinds of substances, not just food. But when the mouse had access to as much food as it wanted, it no longer drank as much alcohol. Based on this information, many alcohol researchers concluded that the mouse was either drinking a lot because it was motivated by hunger or because it was simply behaving abnormally—but in either case it was not a good model for human alcoholism. The researchers who developed the timed model pointed out that their experiments on how much water restriction was needed to induce heavy drinking suggested that thirst was playing only a very small motivational component in their model, but grant reviewers were skeptical about the wisdom of introducing any new motivational components into the experimental setting.

The failure of the timed model in the review panels and ensuing discussions about the merits of both the timed drinking and evening drinking models reveals that "motivation" is one of the contested sites in building the epistemic scaffolding to support research with these models of human binge drinking. Creating a protocol where mice will drink enough to meet the physiological benchmark of a 0.08 % blood alcohol level was not enough to satisfy some alcohol researchers' conceptions of what counts as a valid animal model of alcoholism; these researchers also want to know *why* the animal is drinking so much. Using water deprivation as a motivator to get mice to drink more alcohol contravened a longstanding premise in the alcohol field that "alcohol should be consumed for its pharmacological effects, and not strictly for its caloric value or taste" in animal models for alcoholism (Cicero 1979, as cited in Li, Lumeng, & Doolittle, 1993). This standard

is one of several criteria orginally outlined by Cicero for evaluating whether an animal model is a valid model of human alcoholism that are still widely cited in the field today.[4] The standard that animals must be consuming alcohol for its pharmacological effects specifies what kinds of links need to be made for the epistemic framework of an animal model of alcoholism to be considered solid. It also contains an implicit image of the human drinker. The assumption that underlies this criterion is that human alcoholics drink because they want to experience the pharmacological effects of alcohol, and so too should mouse models of alcoholism in order to create strong links between the animal and the human.

In response to these criticisms, some ARC researchers attempted to counter arguments against the timed drinking model by complicating the representation of the human drinker embedded in Cicero's (1979) criteria for a valid animal model of alcoholism. For the animal modelers in the ARC who supported the timed model, the idea of alcoholism as a complex, heterogeneous disorder provided resources for arguing for the utility and validity of the model. If human alcoholism is a complex disorder with multiple trajectories, then researchers may not need to come to the consensus on a protocols that models "binge drinking" as long as each new model offers a useful representation of some aspect of human binge drinking. In a paper describing the evening drinking protocol, the authors argued that Cicero's criteria were nearly impossible for any one animal model to fulfil. They suggested instead that "a realistic and useful alternative approach is to develop multiple, partial models, each of which addresses a subset of the features." Researchers also argued that the diversity that exists in researchers' animal modeling practices may

[4] The criteria for a valid animal model of alcoholism, based on Cicero's original criteria, are: (1) animals should orally self-administer ethanol; (2) the amount of ethanol consumed should result in pharmacologically relevant blood ethanol levels; (3) ethanol should be consumed for its pharmacological effects, and not strictly for its caloric value or taste; (4) the animals must be willing to work for ethanol; (5) chronic ethanol consumption should lead to tolerance; (6) chronic ethanol consumption should lead to dependence; and (7) an animal model of alcoholism should display characteristics associated with relapse (this summary is adapted from the description availabe on the Indiana Alcohol Research Center's website, `http://medicine.iupui.edu/iarc/animal/`).

even be seen as an advantage, since multiple models with slightly different results might reflect the multiple trajectories that humans take towards alcohol dependence. Raymond notes, for example, that researchers have created several lines of high drinking and low drinking rats through breeding procedures, and each of these lines has slightly different patterns of alcohol consumption. While some rats drink steadily throughout the day, others drink large amounts and then none at all, a pattern that resembles human binge drinking. The many different lines of animals and tests that are available in the alcohol research community, he suggests, might actually be well suited to capturing the complexity of alcoholism.

One way in which ARC researchers challenged existing representations of the human in the epistemic scaffolding of alcohol research was by turning the question away from whether thirst, taste, or desire for pharmacological effects draws mice to the sipper bottle and towards the motivations of the human drinker. Linda, who was one of the early adopters of the timed drinking model and who was also new to the alcohol research field when she first joined the consortium, recalled that she was surprised at the research community's response to the use of fluid deprivation in the timed model. She saw no intrinsic problem with a model that used fluid deprivation, because she thought that human alcoholics might also be motivated to drink because of thirst. She says:

> The fluid deprivation issue is a big no-no in the alcohol field, although alcoholics are clearly fluid deprived. These people do not drink water and hydrate themselves properly. But nevertheless, alcohol folks have established over decades of research that fluid deprivation is bad. So that's when we met, and as a group, and it was decided that the [timed drinking] model was no longer a high priority model, and in fact I was discouraged from submitting a grant using [timed drinking].

Linda complicates representations of human alcoholism in alcohol research by drawing additional information about the human drinker into discussions of model validity. She presents a picture of alcoholism as something that is more complex than just drinking for pharmacological effect, and may be directed by other kinds of motivations. By altering the

representation of human drinking that is embedded in the epistemic scaffolding of alcohol research, she turns features that were widely portrayed as limitations of the timed drinking model into reasons why it might actually be an even more accurate representation of human alcoholics than models where mice drink for pharmacological effects alone.

Researchers also argued that the work that needs to be done to get mice to drink in experimental settings could be thought of as similar to aspects of the trajectory of human drinking behaviors. Consortium members pointed out that the idea that humans are drinking solely for pharmacological effects may be true for some stages of their drinking careers, but not others. Frank, for example, suggests that the idea that mice might be drinking (or not drinking) because of the taste of alcohol, especially initially, is entirely compatible to how humans learn to drink. He says:

> Oh, taste has a huge amount to do with it. Anyone who thinks that taste doesn't have anything to do with it is just fooling themselves. I think Samson would say that taste has to do with it when humans drink…it doesn't really, it's not a problem that taste factors into it. He would say that we all go through a sucrose fading procedure, that when we started to drink we drank these Mike's hard lemonades or whatever.

Frank's argument modifies representations of the human by suggesting that human alcoholics also have to work their way up to being heavy drinkers, just like mice. Frank compares some of the experimental techniques that have been used to get rodents to drink to situations where humans learn how to drink. He cites fellow animal behavior geneticist Herman Samson's work on "sucrose fading," a procedure for getting rodents to drink that involves mixing sugar into the alcohol solution to make it more palatable, and then gradually reducing the amount of sugar over several weeks until they are drinking unadulterated alcohol. Like the timed drinking and the schedule-induced polydypsia models, alcohol researchers voiced concerns about whether the mice in this model were really motivated to drink alcohol or were simply tricked into drinking alcohol when they really wanted the sugar. But Frank points out that even human alcoholics don't start with

straight vodka, and most human drinkers initially drink alcohol in forms that mask its taste. While human alcoholics might eventually drink alcohol primarily for its pharmacological effects, this does not mean that taste is irrelevant at all stages of drinking. Once again, Frank's links between mouse models and human drinkers reverse the discussions about the limitations of mouse models by complicating the image of the human drinker. Frank portrays the procedures that researchers use to "initiate" mice to drinking alcohol in the lab as similar to the way that humans learn to drink alcohol, strengthening the epistemic scaffolds of drinking models.

To summarize, the existing epistemic scaffolding of animal model research portrays the human drinker as one who is motivated to drink by a desire to experience the intoxicating effects of alcohol, and not for other reasons. In their attempts to defend their new models as valid representations of the human, the ARC researchers drew on culturally available ideas about alcoholism as a complex disorder with multiple factors and trajectories that lead to addiction. NIAAA reviewers argued that the timed drinking model in particular was an inadequate representation of human binge drinking behavior by pointing to the multiple motivations that might push mice to drink such large amounts of alcohol, but other researchers argued that the factors motivating mouse drinking were no different than the complex of factors that influence drinking in humans. Researchers offered examples to argue that human drinking is not only about getting drunk, but also about satisfying biological urges such as thirst or using sweetened beverages to overcome humans' own aversion to the taste of alcohol. In this way, ARC researchers modified the representation of the human drinker to include new kinds of motivations for drinking through their attempts to stabilize the shaky epistemic scaffolds of the new models.

3.3 Understanding Human Environments through the Mouse Phenome Database

So far I have argued that the epistemic scaffolding of animal behavior genetics experiments act as conceptual metaphors that researchers can use to highlight or mask particular features of both the mouse and the human. The configuration of these experiments and the scaffolding that supports them makes it easier to understand human or mouse behavior in some ways rather than others. The epistemic scaffolding of the timed drinking and evening drinking protocols makes it easy to see even reluctant drinkers like rodents as comparable to human binge drinkers, at least in some respects. Likewise, researchers hope that the drinking behaviors of genetically controlled animals will make the heritable aspects of human drinking behaviors more visible. I have also argued that the "entailments" of experimental metaphors like the timed drinking and evening drinking models can be flexible, and showed how researchers use the process of getting mice to drink to highlight the environmental contingencies that promote human binge drinking and the complex motivations of human drinkers.

In this section, I will introduce a final example to explore how representations of the human are structured in areas of biomedical research where the human being articulated is more vague. In making binge drinking models, researchers are drawing on existing clinical information and cultural understandings of human binge drinking, and relating this information directly to their models. In other situations, the type of human that researchers aim to represent in their models is less clearly defined. This final section will look at how the "environment" is structured in the Mouse Phenome Database (MPD), a database of mouse information designed to facilitate "translational" research on human health. The construction of this database reveals different disciplinary assumptions implicit in biomedical research with mice about how much power the environment has to alter mouse and human behavior or physiology

Like the status of alcoholism as a disease, beliefs about the importance of the environment in shaping human behavior also have cultural and historical trajectories. In his historical exploration of the impact of Darwinism, Degler (1991) follows the way in which the balance between environmental and biological explanations has shifted over time in American social thought. Degler argues that the idea of a biological root to many human behaviors was nearly universally accepted in scientific and popular thought at the beginning of the twentieth century, only to be nearly universally rejected by sociologists, psychologists, and anthropologists near mid-century. Where once scientists advocated eugenic programs to eradicate undesirable human behaviors, by 1950 the psychologist B. F. Skinner proclaimed that human behavior was nothing more than a conditioned response to environmental stimuli. Degler then traces the way that biological explanations gradually re-emerged in the latter half of the twentieth century in disciplines like psychology and anthropology and in emerging scientific movements like sociobiology.[5]

The question of what kind of environment should be considered relevant for genetic research and how much influence researchers should anticipate it will have is also at play within and between scientific fields. As I explored in the first chapter, the researchers in the Smith laboratory have a particular set of expectations and shared assumptions about what variables constitute "the environment" of the laboratory and how much impact they think particular factors have on mouse behavior. But in experimental reports on behavior genetics results, the environmental contributors to behavior are typically part of the vague and unarticulated backstage of producing experimental results. Mathias, a behaviorist from Germany, observed:

> It's kind of … it's almost schizophrenic in a way, because standardization is based on the realization that the environment affects your results. Because you find that the environment affects your results, you try to standardize it so that it can't affect them anymore. By this of course you basically try to get rid of the environment.

[5] See also Nelkin & Lindee, 1995 on the resurgence of biological (and specifically genetic) explanations for human behavior in American popular culture.

Mathias points out that behavior geneticists' obsessive control of the laboratory environment is based on an appreciation for how much the environment affects behavior, but in controlling for the laboratory environment they make it disappear into the background of the results that they eventually produce. By amassing information on the invisible spaces in the laboratory like mouse housing environments, the Mouse Phenome Project makes these spaces visible again, revealing what kind of environmental factors researchers control form and how they impact research results.

The Mouse Phenome Project is an effort to build a database of information about the "phenome" of the laboratory mouse; that is, measurements on the physical characteristics of different inbred strains of laboratory mice such as body weight, blood pressure, bone density, activity levels, or anxiety scores. The project is organized and administrated by the Jackson Laboratory (JAX), the non-profit corporation that has acted as the central repository for strains of research mice for the last fifty years (Rader, 2004). A group of mouse researchers from various disciplines first discussed the concept of a phenotype database in an informal meeting at JAX in 1999 (Paigen & Eppig, 2000), and the online database was officially introduced in 2004 (Bogue & Grubb, 2004). The database contains information on the baseline anatomy, physiology, and behaviors of 40 different mouse strains, and mouse researchers whose experiments met the submission criteria for the project are encouraged to submit their data for inclusion in the database.

The goal of the project is to provide mouse researchers with a way to compare the phenotypic characteristics of each strain so that they can choose mouse strains for new drug discovery and toxicology studies, or for developing new models for human diseases. By merging extensive data about mouse genomics with new information about the physical characteristics of different mouse strains, the project developers hope that researchers will be able to identify new correlations and make predictions about the role of genes in particular diseases. For example, in the original proposal describing the concept of the

Mouse Phenome Project, the authors pointed to studies on salt-induced hypertension using rat and mouse models that located many of the same genomic regions as human studies, but in one-fifth of the time and for much less money (Paigen & Eppig, 2000, p. 715).

The MPD was not only envisioned as a new epistemic and experimental structure to make genes visible, but also to make environmental factors visible. The MPD developers hoped that amassing a large amount of baseline data about this group of mouse strains would provide bioinformaticians with material to analyze gene-environment interactions. When I interviewed Molly Bogue, the director of the project, in 2008, she described how the dataset could be used in this fashion:

> The other thing that's happening now is statisticians and computational people are coming in, and really they are taking just about everything that they can get their hands on, and doing this huge data crunch stuff. And that's really exciting, because that's where the really, you know, completely unpredictable things are going to start falling out. I mean you can predict things with this complex kind of systems biology approach that there's no way you could do on a gene-by-gene basis. So that I think is extremely exciting.

In some ways, this approach can be thought of as an alternative way to solve the problem of variation in researchers' experimental practices that the founders of INIA described. Instead of attempting to direct researchers' attention towards a single model, the Mouse Phenome Database turns the existing variation in experimental protocols into a dataset that can be used to investigate what factors influence mouse phenotypes. In order for the database to function in this way, the project organizers argued that detailed information on the source of mice, their diet, their environment, and their general health would be needed (Bogue & Grubb, 2004, p. 73). Along with each data set, the project curators also collect detailed information about the protocols used to generate the measurements and the housing environments of the mice.

As Shostak (2007) has observed, the laboratory mouse is a research object that is used so widely in contemporary biomedical research that it makes links across many social worlds.

She argues that the special nature of genetically modified mice, which are "simultaneously molecular tools and whole animals," has facilitated communication and cooperation between researchers with vastly different disciplinary backgrounds and methodological approaches, such as toxicologists and molecular biologists (p. 321). The MPD provides a similar space where researchers from many different backgrounds can interact through inbred strains of mice as common objects. While the ARC consortium aimed to bring together an interdisciplinary group of alcohol researchers with behavioral, molecular, and neuroanatomical backgrounds, the researchers that the MPD brings into contact with each other are even more wide-ranging: The database includes contributions from fields as diverse as behavior genetics, toxicology, and cancer research, as well as researchers working in commercial drug development settings. Developing and implementing standards of reporting to span these different communities of practice reveals that there are substantially different visions of the environment implicit in different mouse research communities.

The project developers settled on what they thought was a "fairly minimal" set of reporting requirements about researchers' test environments and housing conditions, such as the type of bedding used and the temperature of the colony rooms. Bogue recalls that she thought these variables were "no brainers" to include in the database, because in her view it was widely known that these types of factors vary between laboratories and can affect experimental outcomes. But when researchers began submitting protocols, Bogue recalls that not all of the researchers submitted information on the variables that she thought were necessary to report on. She says:

> If we don't have the information that we think we need, we keep going back to the investigator. Which, you know, that can have problems too. The ones that don't really appreciate this detail or the need for the detail are kind of like, what is these people's problem? [laughter] Why do they keep bugging us, you know? Is it important that vitamins have been supplemented in our water? Is it really that important? Well, yeah, we need to record that.

In some cases when the MPD curators contacted researchers who had submitted data for inclusion in the database, they found that researchers didn't include certain pieces of

information in their submissions because they didn't record factors like the temperatures of their colony rooms. On the other hand, some researchers submitted datasets with more metadata than even the MPD curators had anticipated. Although the project has a policy of always including whatever information the submitting investigators think is relevant to the experiment in the database, some information seems to fall outside the norms of what they think is necessary and reasonable to report on. For example, when I visited the MPD, I asked one of the project curators to show me a sample data set and how the protocol information and metadata would be recorded in the database. She showed me a data set from a group of alcohol researchers on a two bottle choice experiment. She says:

> It's a small [protocol], that's all there is to it, but there's so many steps to it. In two days, the tubes are changed this way to avoid the mice being fixed on one bottle compared to the other, and how close [the bottle] is to the food … I mean, those behaviorists, they are so anal! [laughter] Always shuffling things, I don't know, if I was a mouse, I'd be like "wait a minute, I was used to drinking the water on this side! And now the water's on this side!"

Behaviorists may be at the extreme in the mouse community for what they think counts as an environmental variable, but even within the behavior genetics community, there is a substantial amount of variation in the way that researchers control and report on aspects of the experimental system such as housing environments for their mice. For example, many articles in one of the field's main journals, *Behavior Genetics*, employ similar boilerplate descriptions of the mouse housing environment, such as "mouse colony rooms were maintained on a 12/12 light/dark cycle," "two to five same-sex littermates were housed per cage," or "all mice were maintained with food and water ad libitum" that offer standard pieces of information about the laboratory environment, but some articles report on many more variables and in much greater detail. An excerpt from a more extreme example of a methods section from one paper demonstrates the level of detail that some researchers think is necessary to report in the scientific literature:

> The animal room was sound-attenuated. Relative humidity was kept at a constant level of 50 %, the ambient temperature was maintained at 21.0±2.0 ï¿½C

> and the ventilation rate was 15–20 air changes per hour. To reduce stress in the laboratory animal facility, during the whole day (24 h) radio-sound (60±3 dB) was provided. In addition conversational radio-sound (e. g. talk shows) may accustom the animal to the human voice … All mice were housed individually directly after arrival in enriched, wire topped Macrolonï¿½ Type II L (prolonged) cages (size: 365×207×140 mm, floor area 530 cm²; Techniplast, Milan, Italy). Enrichment, besides standard bedding material, included a shelter, a tissue (Kleenexï¿½: Kimberly-Clark Professional BV, Ede, The Netherlands) and a small amount (less than a hand full) of paper shreds (EnviroDriï¿½: Technilab-BMI BV, Someren, The Netherlands). (Laarakker, Ohl, & Lith, 2008, p. 160–161).

This excerpt shows both the specificity (down to the brand of cages and the sound pressure) and the range of factors (such as humidity and ventilation rate) that some animal behavior genetics researchers report on in the literature.

As I have discussed previously, this assessment of what counts in the laboratory environment has a disciplinary aspect to it, and groups of researchers also develop local standards about what counts as "the same" or "well-controlled" in their experimental practice. But the types of environmental variables that researchers record and report on are also tied to visions of the human. Descriptions about the importance of collecting information on the laboratory environment also contain implicit assumptions about what kinds of factors are (or could be) relevant in human environments. For example, Bogue describes the importance of collecting metadata on animal housing conditions as follows:

> And they, so they can go back and say "oh, this is clustering with this under these conditions but not those conditions. What's different? Oh, this is a high fat diet, and this is a low fat diet. Oh, wow! So this gene pathway is involved in blah blah blah, and when there's not a high fat diet you don't see that." So on a really large scale analysis point of view, the detailed protocols are really needed for that.

This example of the "high fat diet" and how it might be relevant to understanding gene action is one that appears several times in descriptions about the potential benefits of the Mouse Phenome Database (see also Paigen & Eppig, 2000, p. 717). This description of the laboratory environment is one that I suspect is a frequently used example because

of its treatment relevant entailments: It readily lends itself to speculations about how a factor that is important in the mouse's environment could also be relevant in a human clinical setting. If mouse researchers were successful in identifying a gene whose action is modified in response to changes in diet, then doctors could advise their human patients to adopt a low-fat diet to mitigate the consequences of a risky allele. This example aligns environmental factors that are relevant in experimental mouse worlds with factors that are also relevant in human worlds to show the potential translational benefits of mouse biomedical research.

In other cases, collecting metadata on researchers' protocols in projects like the Mouse Phenome Database might reveal information that is not necessarily what researchers had in mind. For example, one behavioral researcher who did an analysis of the metadata he recorded in his own laboratory found that the single biggest variable affecting the results of his pain tests was the person who performed the experiment (Chesler, Wilson, Lariviere, Rodriguez-Zas, & Mogil, 2002). The mouse's genotype and the season and time of day of testing accounted for part of the mouse's latency to withdraw its tail from a hot plate in a pain test, but the researchers found that the biggest factor determining a mouse's score on this test was the experimenter. In a letter to the editor in *Nature Neuroscience,* the authors of the study cautioned that projects like the Mouse Phenome Database might encounter similar problems in analyzing metadata:

> Large projects are often carried out by multiple undergraduate and graduate students, postdocs, technicians and other transient personnel. Furthermore, research, particularly on mutant mice, increasingly involves collaborations. The impact of these "nuisance" factors becomes greater as data are shared in the growing body of online resources, such as those generated by large-scale phenotyping efforts, including mutagenesis screens and the Mouse Phenome Project. (Chesler et al., 2002, p. 1102)

The authors' description of the experimenter effect as a "nuisance variable" strongly suggests that this is not the type of environmental factor that researchers have in mind when they are looking for gene-environment interactions. Despret (2004) has argued that re-

searchers in American psychology have often focused on ways to eliminate the relational aspects of animal research rather than treating it as interesting information. She points to Rosenthal's work on experimenter effects, which she argues draws attention to the effects of the experimenter on the experimental subject for the sole purpose of figuring out how to erase the experimenter from the experimental setting. In her words, "Rosenthal's original idea had not been to explore a world enriched and created by these differences; it had been to mark them off as parasitic supplements that seriously contaminate the purity of the experiment" (p. 118). Similarly, the relational aspects of Chesler et al.'s (2002) findings are treated as unwelcome interruptions to the experimental setting rather than important sources of information about pain. In making the laboratory environment in mouse research visible, the Mouse Phenome Project also makes visible a few features about how much capacity different researchers think the environment has to shape behavior and what kind of environment is envisioned for humans through mouse biomedical research. The environment invoked in descriptions of the Mouse Phenome Database is a treatment-relevant environment that individuated, not relational.

3.4 Conclusion

This chapter showed how representations of the human are articulated through the process of building epistemic scaffolds to support models of human behaviors. Although animal behavior genetics researchers consider themselves to be "basic" scientists, their work with animal models also has an applied focus that continually references back to human health. Even in projects such as the Mouse Phenome Project that seem to be quite far away from applications to the human world, there are still implicit images of the human embedded in mouse research. To explore how the human is represented in animal models, I treated animal model systems and their associated epistemic scaffolds as metaphors that allow researchers to understand one organism in terms of the other. As metaphorical concepts,

animal models have particular entailments that make some aspects of human behaviors especially visible and mask others.

I also explored some of the many representations of the human that emerge through the process of animal research: the treatable human, whose genes and diet could be modified for better health; the complex drinker, who drinks alcohol because of a combination of different motivations and aversions; and the environmentally controlled drinker, who drinks more or less depending on the structural features of his or her surroundings. In contrast to criticisms of behavior genetics research that suggest that it promotes genetically reductionist represenations of the human, I have shown some of the ways in which practitioners use experiments as metaphors to highlight the importance of environmental features or other aspects of human biology. Animal behavior geneticists themselves also describe their method as reductionist, albeit in a more qualified fashion. While they argue that human behavioral disorders like alcoholism have substantial genetic and biological components, they also argue that not all aspects of alcoholism can be reduced to genes.

While researchers may view the information produced by mouse research in the laboratory as a set of pragmatically crafted partial truths about the reality of human binge drinking, they are nonetheless potentially consequential representations of alcoholism. Campbell (2000, 2007), for example, argues that cultural representations of drug users shape the way that public policies for addressing drug use are developed. Although the complex process of stabilizing behaviors in the laboratory may not be forgotten by researchers, environmental contributors to behavior may not always be visible in descriptions of the findings that these experimental systems produce. Cwiartka (forthcoming) argues that data from animal experiments may be especially problematic in new contexts such as policy settings, because the nuances of the degree of relatedness between the animal and the human tend to disappear once the link between the two is forged. The flexible uses of animal experimental systems as metaphors for the human that I described

here may reflect the fact that many of the researchers I interviewed were heavily involved in the methodological aspects of developing these new models. The entailments of these experimental systems may not be so flexible for all users of the models or in all settings, such as when the behavior genetics facts produced with these models are taken up in new knowledge communities.

4 The Mouse Multiple? The Career of the Mouse in the Laboratory

After an early morning seminar series, I walked with Grace back towards the Smith laboratory offices. She had presented some of her data in this session on stress and drinking behavior in young mice, and was frustrated with the response to her talk. The scientific literature suggested that stressful experiences should impact drinking behavior in adolescent rodents, but she hadn't been able to find any effect in her young mice. The audience hadn't seemed prepared to accept this conclusion at face value, however, and most of the questions after her talk focused on the types of procedures she had used in her experiment and whether they counted as "stressful." Grace recalled the comments of one audience member, who said that he had produced good results in his laboratory by using different kinds of stressors at unpredictable times of the day. "The problem with stress," Grace told me, "is that everyone's done it at some point in their careers." The probability of getting these kinds of methodological questions in a talk was pretty high, she said, because everyone thought that they were knowledgeable enough to say something about it.

On another early morning in a windowless conference room, I observed quietly at the side of the room as members of the Institutional Animal Care and Use Committee (IACUC) flipped through binders of information on protocols that the committee would be reviewing on that day. One member paused at a description of a behavioral experiment, the "tail suspension test." The protocol description explained that in this test, a mouse is held upside down by its tail while the researcher measures how much time the mouse

spends moving and how much time it spends passively hanging. Researchers use the test to model depression, and anti-depressant drugs will increase the amount of time that the mouse spends struggling to escape before giving up and simply hanging. "Is a 15 minute suspension normal?" the committee member asked. Laura, also a member of the committee, replied that it was. "And is the mouse in pain?" the member asked again. Laura paused. "It's not exactly painful," she said, "but the mouse is in some distress."

In both of these instances someone is assessing whether a mouse is stressed (or in distress), but the circumstances and the people involved are very different. Grace and her colleagues are concerned with "stress" as an experimental variable, something that will alter a mouse's behavior and biology. The IACUC committee member is concerned with the mouse's subjective experience of the proposed experiment and how much "distress" the animal will feel. Is "stress" the same thing in each one of these situations? Or, to put the question slightly differently: Is the stressed-out mouse in each of these situations the same mouse?

So far in this dissertation, I have focused mainly on the concerns and actions of researchers in the Smith laboratory as they are building experimental systems and epistemic scaffolds, but their concerns about how to produce good experimental data using mice are not the only considerations directing action in the animal behavior genetics laboratory. Experimental work is shaped by many sets of actors with different goals and interests: While scientists might be focused on producing good experimental data, veterinarians may be worried about stopping the spread of mouse diseases, commercial mouse producers may be concerned with protecting their intellectual property, and institutional regulators might be focused on minimizing the pain and suffering of animals used in research. Creating scientific knowledge involves more than just the construction of theoretical concepts and technical solutions; it also requires the institutional, financial, and political support that makes particular research projects "doable." (Fujimura, 1996)

This chapter returns again to the Smith laboratory to take a broader view of animal behavior geneticists' knowledge production practices. I lay out the "career" of the laboratory mouse as it moves through the laboratory on its way to becoming acceptable experimental data as a way of revealing some of the institutional settings and social worlds in which animal behavior genetics research is embedded. In following the mouse from birth to death, I examine different moments where humans come into contact with the mouse and how the interactions between different actors and regulations plays out in the shared space of the laboratory. I ask: Who interacts with the mouse, and what tools and techniques do they use? What status do they assign to the mouse, and what kind of ideal mouse are they trying to realize? How do the Smith laboratory researchers talk about the intersections or separations between these factors?

In particular, I focus on the relationship between the practices that constitute the mouse as an experimental subject and the mouse as an ethical subject—practices that intersect in some moments in the animal behavior genetics laboratory. As the introductory examples show, some concepts like "stress" that are embedded in animal behavior geneticists' epistemic scaffolds share semantic connections with concepts such as "distress" that are used in the welfare world. Welfare practitioners and animal behavior geneticists also share some tools and practices. Mouse huts, mouse handling, reducing ambient noise, and using standardized cages are all practices that could belong to two different systems of management and meaning in the behavior genetics laboratory. At times, researchers even find themselves tasked with the management of both the mouse as ethical subject and the mouse as experimental subject. Scientists design experimental protocols to investigate stress, but also adjudicate on distress in IACUC meetings. Technicians control for stress during experiments, but also worry about the distress they are causing to other living beings. I explore some of the ways in which researchers in the Smith laboratory conceptualize the difference and sameness of these concepts and practices, and (if necessary)

work to resolve ambiguities created by differences. Researchers talk about some institutional structures for regulating animal welfare, such as the IACUC, as well integrated with existing experimental practices and providing ethical oversight and justification for mouse research. In other cases, researchers talk about welfare practitioners as though they are participating in or using animal behavior geneticists' epistemic scaffolds, creating ambiguities in experimental work. In either case, looking at the various regulations and actors that act on the body of the mouse offers a way of thinking about how animal behavior geneticists' epistemic scaffolds and epistemic cultures are shaped by particular institutional locations.

4.1 The Career of the Laboratory Mouse

To organize my discussion of the transit of the mouse through the Smith laboratory, I employ Goffman's (1961) the concept of the "career" pathway. In his discussion of mental hospitals, Goffman argues that the gradual transition of a person into a mental patient can be conceptualized as a "moral career" that has a regular series of stages that the patient goes through, marked by transitions in the way that the person views the world and his standing in it. He emphasizes both the commonalities of the stages that most mental patients go through, such as bureaucratic processes and adapting to new institutional settings, as well as "contingencies" that can unpredictably shape a mental patient's career, such as whether or not the patient finds himself in close proximity to a mental hospital during a crucial moment, or instances where an alcoholic might be sent to a mental hospital because a jail is full. Glaser and Strauss (1968) developed a similar concept of an "illness trajectory" that describes a series of phases that patients pass through in the hospital, and how the status or evaluation of the patient changes as they pass through this trajectory. Lynch, Cole, and McNally (2008) have also used Goffman's concept of the career to describe the transit and transformation of inanimate objects, such as how crime scene samples pass through a

series of stages on the way to becoming criminal evidence that changes their moral and epistemic status as "good" or "bad" evidence.

The mouse's transit through the laboratory can also be thought of as a "career" that has a regular series of stages that the mouse passes through on its way to obtaining its status as good experimental data. This concept is useful for tracing out some of the many different actors, practices, tools, and institutional processes that are acting on the laboratory mouse. Goffman argues that the concept of the "career" is is well-suited for studying the interactions between individuals and institutions because of the dual meaning of "career." He writes:

> One side is linked to internal matters held dearly and closely, such as image of self and felt identity; the other side concerns official position, jural relations, and style of life, and is part of a publicly accessible institutional complex. The concept of career, then, allows one to move back and forth between the personal and the public, between the self and its significant society, without having to rely overly for data upon what the person says he thinks he imagines himself to be. (Goffman, 1961, p. 127)

This chapter is not concerned with transformations in a mouse's own identity (if it can be said to have one), but transformations in the ways in which different actors view the mouse and the characteristics that an ideal laboratory mouse should have at different points at time. By tracking the body of the mouse through the laboratory and focusing on moments where it interacts with different people and regulations, it is possible to examine the interplay between actors' assessments of what makes a "good" mouse (experimental or otherwise) and the institutional structures that the laboratory is embedded in.

Examining moments where different social groups or regulations act on the body of the mouse also provides opportunites to explore the relationships and interactions between different practices, actors, and statuses that they assign to the mouse. In his study of research practice in the neurosciences, for example, Lynch (1988) distinguishes between two different kinds of animals in the laboratory: the alive, fuzzy, "naturalistic" animal that has perceptions and subjective experiences, and the "analytical" animal that is an artifact

of human intervention. He argues that the production of scientific data in the laboratory tacitly depends on "commonsense" knowledge about animals, such as knowing how to breed and handle animals in the laboratory. "Like physicians who show good 'bedside manner'," Lynch writes, "scientists who are 'good with animals' can sometimes obtain compliance from their subjects which otherwise would be impossible" (p. 280). He focuses on the way that scientists use these practices to mediate the animal's transition from a holistic animal into acceptable experimental data. Lynch describes these skills as a kind of "subjugated knowledge" (p. 267) that is necessary for producing scientific knowledge, but is not recognized as an important form of knowledge in and of itself.

In her ethnographic study of how atherosclerosis is treated in the clinic, Mol (2002) takes the idea of differentiating between different practices and conceptions of a single object much further. She dissects the seemingly unitary disease of "atherosclerosis" by paying close attention to how different actors are "doing" or "enacting" the disease differently. For the pathologist with a microscope, atherosclerosis is a disease of thickened or blocked arteries; but for the physician talking to a patient in the clinic, atherosclerosis is a disease that causes pain and limits patients' mobility. Mol argues that these ways of differently enacting the disease are not just different representations or understandings of the disease, but can be conceptualized as different diseases. Atherosclerosis, and by implication the body that the disease resides in, is "multiple," in her words. She argues that doctors, patients, researchers, surgeons, physical therapists and diagnostic technicians each bring a different "atherosclerosis" into being through their instruments and practices.

Mol suggests that very different assessments of what "atherosclerosis" is may seem to imply fragmentation and conflict, but she argues that these different atherosclerоses can still "hang together" in the body of the patient. She describes several processes that allow multiple versions of the disease to co-exist in the same patient and the same clinic. First, she argues that even though patients, doctors, pathologists, and surgeons may be enacting

different versions of the disease, these multiple enactments of the disease often "cohere" together. There is enough correspondence between the versions of the disease that doctors and patients can jointly enact atherosclerosis in the therapeutic encounter. Second, she argues that tensions and frictions between different versions of the disease rarely erupt into controversy because these tensions are "distributed" through the spaces of the clinic and the course of treatment. General practitioners and surgeons may be working with different versions of the disease, but tensions between these two approaches are diffused because they are enacted in different locations in the clinic and at different times in the course of the disease. These processes of coordination and distribution allow doctors, clinicians, and patients to find ways to negotiate and live with the incompatibilities of different ways of enacting the disease without ever coming to agreement on what atherosclerosis really is.

Mol's "praxiographic" analysis of the clinic emphasizes the differences between the ways that actors enact shared objects; but another way of conceptualizing the relationship between different objects and concepts in the laboratory is to think of them as mutually constitutive, or "co-produced." Jasanoff (2004) argues that analysts should not presume that there is a fundamental separation between categories and concepts that we ascribe to the "natural" world and those that we describe as social categories. As she puts it, the co-productionist argument suggests that "we gain explanatory power by thinking of natural and social orders as being produced together … Knowledge and its material embodiments are at once products of social work and constitutive of forms of social life; society cannot function without knowledge any more than knowledge can exist without appropriate social supports." (p. 2–3)

In her study of the Human Genome Diversity Project (HGDP), Reardon (2001, 2005) uses the concept of co-production to describe the interaction of science and ethics. In some instances, she uses it to describe the concomitant development of scientific knowledge with

social practices for governing that knowledge. As she puts it, co-production demonstrates the "ways in which ordering the natural world requires not just scientific ideas and practices, but also laws, norms of ethical practice, and credible governing arrangements" (Reardon, 2001, p. 358). She points to instances like the HGDP's ultimately unsuccessful attempts to develop new ethics protocols that would provide a mechanism for groups to consent to be studied alongside new experimental practices for sampling groups. Reardon also uses co-production to describe the interpenetration of scientific and social categories. The scientific meaning of the term "race," she argues, is bound up with its meaning in other contexts, and therefore the HGDP organizers could not reasonably expect to generate a new scientific definition of race without raising concerns about what this definition would mean in other locations. She writes:

> Rather than a category that respects any demarcation between the scientific and the social realm, race has traveled vigorously and often across the boundaries of science and society, reality and ideology, throughout the twentieth century. In the the process it has been stabilized and destabilized, made and remade. (Reardon, 2005, p. 9)

She points to several instances of transit between the scientific and social meanings of race, such as the researchers' circular attempts to define biological groups for study using anthropological data and fears from HGDP opponents that information about biological relatedness would rewrite social understandings.

Each of these analytical positions offers different ways of thinking about the interactions of actors and their understandings of the mouse in the Smith laboratory, placing emphasis on different moments in the career of the mouse: transit between understandings; practices, tools, and spaces that constitute different understandings; and the interplay and mutually-reinforcing qualities of different understandings. In the remainder of this chapter, I draw on my fieldwork in the Smith laboratory to describe the career trajectory of a fictional mouse that highlights the variety of actors, tools, and practices acting on the mouse in the animal behavior genetics laboratory. Using some of the conceptual resources outlined

above, I also examine the ways in which the Smith laboratory researchers describe the interaction (or lack therof) of different practices and statuses assigned to the mouse in the laboratory.

4.2 Before the Experiment

Day 1: Mouse X begins its journey through the laboratory in a hard plastic cage about the size of a shoe box. The bottom of the cage is covered with bedding made of recycled paper that looks like small grey pebbles, and a mass of shredded paper resembling a clump of cotton balls occupies one corner. A pile of bright pink baby mice and a female adult mouse are sleeping in an indentation in the paper fluff, while a male mouse scratches in another corner. At just a few hours old, Mouse X and his siblings are hairless and blind, with their eyelids closed and their eyes just visible underneath their translucent skin, and they stay close to their parents and the shredded paper nest for warmth. Although Mouse X can't yet see anything of his surroundings, he may be able to smell that he is not alone in the room. Above, below, and to the sides are other identical cages holding more mice. The "colony room" where Mouse X lives resembles a library of mouse cages: Along the walls and in aisles down the center of the room are shelves holding rows of cages in a grid that extends from wall to wall, and from the floor almost to the ceiling. With an average of three to five mice per cage this space can house thousands of mice at maximum capacity, and is only one of several rooms on this floor that is dedicated to mouse housing.

Mouse X's cage shakes and moves, and the florescent light strips on the ceiling brighten the cage interior as it is pulled out of the rack. It is early in the morning at Western and Rachel, a technician from the Tremblay laboratory, is making rounds through the colony rooms to check on her mice. She sets the cage down on a cart, lifts off the lid that holds a water bottle and a handful of dry mouse chow. She smiles affectionately at the newborn "pinkie" mice that she finds. Without touching them, she counts the pups—nine

in total—and marks the number of new mice and the birth date on a card affixed to the front of the cage. Rachel replaces the cage top and slides Mouse X's cage back into its slot. Moving down the column of cages, Rachel pulls them out like drawers, one by one, to check in on other pairs of breeding mice. When she opens one cage, she is extremely pleased to find a small litter of four newborn pups. "I'm so proud of you!" she exclaims. These mice are a congenic strain bred from two different strains of inbred mice for an experiment in the Tremblay laboratory. Their unique genetic configuration is useful for identifying regions on a chromosome where the Tremblay laboratory hopes to find a candidate gene, but it also seems to have negative side effects on their ability to reproduce. She returns the cage to the rack carefully and wonders aloud if there is anything else she can do to help them along since these mice are already over a hundred days old, quite old to still be breeding.

In the initial stages of his career in the laboratory, Mouse X is regarded as part of the laboratory's "stock" of available mice that may or may not be used for future experiments. Rachel's careful attentions to Mouse X are aimed at producing a *breeding mouse* that can survive and reproduce in the setting of the laboratory. Although not all animal behavior genetics laboratories breed their own mice, reproduction is an important part of the work that takes place in the research facility at Western. It is an essential part of some experimental protocols, such as selective breeding experiments where mice are scored on behavioral tests and then paired based on their scores to create a new generation of mice, bred from high performing or low performing parents. These experiments require substantial facilities to house the large numbers of mice needed (sometimes several hundred mice per selection experiment), and to house breeding pairs of these special lines after selection so that they can be used for future experiments. In other cases, researchers at Western breed mice because they are difficult to obtain from commercial

suppliers, such as some kinds of knockout mice, or simply as a cost-saving measure.[1] Mouse X belongs to a strain that researchers refer to as "het" (heterogeneous) stock: He is descended from a line of mice that were originally bred as the control group for a selection experiment but are now kept in the laboratory for any type of experiment that requires mice with a varied genetic background.

Mouse breeding is a practice that is partially codified in laboratory procedures and partially an art that remains in the realm of lab lore. While heterogeneous strains like Mouse X generally breed readily in a laboratory setting, not all mice reproduce with such ease. Laboratory mice might be too old, too fat, or too sickly to mate, and in many instances the genetic modifications that researchers make in mice for experimental purposes also impair breeding. To increase reproduction rates, researchers typically give breeding mice special bedding materials to build nests and a high-fat food mix, known in the lab as "love chow." In a "desperation move" to try to get the congenic mice to breed, Rachel has also grouped the mice in "harems" of two females and one male. In addition to these practices researchers also talk about the elusive "special touch" attributed to some technicians who are able to get higher breeding rates than anyone else, acknowledging the specialized but difficult to communicate knowledge that successful practitioners of artisanal lab practices hold (Cambrosio & Keating, 1988). Delores, a technician from the Smith laboratory, is particularly skilled in getting mice to breed, and she is responsible for maintaining some of the laboratory's most valuable colonies of mice. Sharon told me that she would like to find some way to capture this breeding knowledge in case Delores ever decided to leave the laboratory, but she found that it was difficult to write down what exactly Delores did that made her a successful breeder. Delores's description of her protocol for pairing, feeding, and housing the mice seemed insufficient to explain the reasons for her unusually high

[1] Laboratory mice today are produced almost exclusively in laboratory settings or commercial production facilities, but other spaces have also historically been important for laboratory mouse breeding. Some of the most commonly used inbred mouse strains, such as C57s and DBAs, originated from lines of "fancy" mice bred by hobby mouse breeders (Rader, 1998, 2004).

reproduction rates.

As Sharon describes it, the way that technicians themselves articulate the reasons for breeding successes are often not compatible with procedural descriptions of how work is done in the lab. As Mol (2002) suggests, actors may try to resolve differences between descriptions and categories, but these attempts at "coordination" are often imperfect. Alexis, a technician in the Martin laboratory who also had a reputation for having a special way with mice, described her ability to get mice to breed in terms of her mood. She told me that she thinks her mice are responsive to her mental state, and so she always tries to take a minute to get into a "calm frame of mind" before going into her colony rooms so that the mice will not sense her being hurried or distracted. Dr. Martin acknowledges the differences in the way that mice respond to members of her staff, but is often unsure of how to incorporate this information into her laboratory's breeding practices. Animal behavior geneticists who take seriously the idea that their technicians' moods may affect their success in breeding mice may translate these descriptions into discussions of pheromones or differences in researchers' smells, but even these biologized descriptions are still difficult to control for in laboratory work.[2]

Day 21: Mouse X and his siblings are growing quickly. At three weeks old their eyes are open, they have a full coat of fur, and they run energetically around the cage. The young mice are such adept jumpers that the technician who comes in to check on them this morning takes care to lift the lid very slowly to prevent the surprised mice from hopping right out of the cage. These mice are ready to be weaned from their mother and moved into a cage of their own. The technician takes Mouse X's cage, places it on a cart, and rolls it over to a workstation where a stack of clean cages with fresh bedding is waiting. He grabs Mouse X by the tail and lifts him into a new cage, along with three other male mice

[2] See, for example, Chesler et al., 2002, where the authors suggest that each researcher's distinctive smell might be responsible for the differences that they observed in data collected by different experimenters.

from his litter. The remaining five female mice are placed into another cage, and mouse X's parents get a clean cage of their own as well. The technician takes the card from the old cage and carefully transcribes the age, sex, and number of mice onto three new cards. He slots these cards into a holder on the front of each cage and puts them on a cart to be rolled back to a slot on the wall.

The gloved hand moving the mice into new cages this morning does not belong to Rachel, but to a technician from the animal care staff. While Rachel has been making her morning rounds to check in on Mouse X over the past three weeks, other technicians and veterinarians have also been visiting to refill the cage lid with scoops of dry mouse chow, replace the water bottle, and make daily rounds to check for mice that look sickly or have signs of infection. Producing an ideal breeding mouse requires the cooperation of these two different workforces: Laboratory workers such as technicians and graduate students that report to a principal investigator, and a separate staff of animal care workers who are not associated with any particular laboratory and are responsible for maintaining the centralized animal housing facilities. Researchers in the Smith laboratory often don't talk much to or about this separate workforce, and the animal care staff operate largely independently. The animal care workers share a bathroom and use the same workstations in the colony rooms as the laboratory staff, but they have their own lunchroom and report to the veterinary staff rather than the principal investigators of a lab. Rachel has arranged her schedule so that she does any work that requires a workstation in the colony rooms in the afternoon, since the animal care tech often use the stations in the morning. Rachel tells me that she knows the names of some of the female animal care technicians, but only because there are so few of them. These routines and divisions of tasks and space allow the animal care workforce and the laboratory workforce to operate as "worlds apart" even though they are caring for the same animal body (Mol, 2002, p. 111).

The division between the animal care staff and the laboratory staff should not be over-

stated, however, because a good deal of coordination is involved in maintaining the independent operation of the two workforces. For many routine matters, Rachel communicates to the animal care technicians by leaving notes on the cages. Housing policies, for example, dictate that only one breeding female should be housed with a male mouse to prevent overcrowding when a litter is born. To prevent her harems of congenic mice from being rearranged by the animal care staff, Rachel leaves instructions not to separate them on the cards on the front of the cages. Finding ways to coordinate is especially important at Western where care of the breeding mouse frequently intersects with the management of the experimental mouse, since researchers often want to maintain control over the tasks that the animal care staff is responsible for. Housing in particular is an ambiguous space in the animal behavior genetics laboratory, somewhere between research and infrastructure, between science and animal welfare. Veterinarians and the animal care staff are primarily responsible for animals who reside in this space, but researchers are also concerned with the management of this space because they see it as an important part of their experimental system. The staff at Western University have instituted unique solutions to negotiate some instances of the conflicting demands of laboratory and animal care work, such as the "do not change" rack that can be found in each colony room on the fifth floor of the Smith laboratory building. Animal care technicians are responsible for changing the bedding in the mouse cages on a weekly schedule, but since many researchers at Western were concerned that fresh bedding too close to a test day might change the behavior of their mice, they designated special racks for cages that technicians should not touch. By moving mouse cages to this rack, researchers indicate that they are taking over this aspect of the animal care technician's role and are now responsible for giving the mice fresh bedding.

In some cases, an ideal experimental mouse and an ideal breeding mouse cannot be produced at the same time. The goals and management strategies of researchers cannot

be fully resolved with those of animal care workers, resulting in conflict. On a visit to the laboratory of David, a researcher at a Canadian university, he took special pride in showing me a small colony room adjoining his laboratory space that he specifically requested when the building was designed. This arrangement was ideal for producing an experimental mouse that would not be stressed by a daily trip between the central housing facility and his lab, but far from ideal for other purposes. He commented that veterinarians hated his colony rooms because they had to travel further to check up on these mice, and they argued that the mice could be better monitored and cared for if they were kept in the central facility. This instance shows how rearranging the typical spatial configuration of laboratory and colony facility also rearranged the typical division of labor between researchers and animal care staff. With the mice located inside the laboratory, it was not clear whether it was the job of the veterinarians to care for the mice, as they normally would for mice who are not actively engaged in experiments, or whether it was the task of the laboratory technicians and grad students.

4.3 Transitioning into the Experiment

Day 42: While Mouse X sleeps in his cage, a virtual proxy of Mouse X pops up on Rachel's computer screen in one of the Tremblay laboratory rooms. Grace, a postdoc from the Smith laboratory, is planning to run an experiment in a few weeks, and she is meeting with Rachel to find out if Rachel has any available mice of the right age and genotype. Rachel pulls up her breeding records, stored in a spreadsheet on her computer. At 42 days old, Mouse X and his siblings are getting close to the ideal age range for Grace's experiments, between 60 and 90 days old. Breeding in the heterogeneous colonies has been going well over the past few months, and Rachel calculates that she can give Grace about ten cages of mice while still leaving enough for the Tremblay laboratory experiments. She writes a list of cage numbers on a slip of paper and hands it to Grace. "Assigning"

Mouse X to an experiment is one of the moments in the career of the laboratory mouse that marks his gradual transition from mouse "stock" into an experimental subject. Now that Mouse X has been assigned to her experiment, Grace will take over Rachel's role as the primary caretaker for Mouse X for the rest of his career in the laboratory.

Like Mouse X, many mice at Western transition into experiments by being assigned from mouse stock maintained in the colony rooms to particular experiments, but this is not the only way that mice are enrolled in experiments. Some mice may be purchased from commerical suppliers such as the Jackson Laboratory or Charles River, and other mice may be obtained from other laboratories specifically for particular experiments. Even the seemingly informal exchange that takes place between Rachel and Grace hints at another way that the mouse is being managed in the laboratory: as a *commercial object* that can be owned, traded, bought, and sold by publicly funded researchers and private corporations. Treating mice and mouse knowledge as valuable commercial entities has become increasingly common in academic research over the past three decades, but researchers at Western actively resist the idea that this is a productive way to regulate the flow of resources in the community. During one interview, Laura redirected one of my questions about technical trends in the field, telling me that one of the developments that she thought deserved more attention (and criticism) was the increasing focus on commercializing genetic knowledge. She said that while most of the researchers at Western had not involved themselves with patenting, she thought that there was increased pressure on animal behavior geneticists to use research advances to create revenue for their laboratories or institutions. A graduate student put it even more tersely: When I asked her about commercial practices in the lab, she asserted that patenting was a trend in *human* genetics, not animal research. In animal behavior genetics, she said, no "self respecting scientist" would patent their mouse models, and those who did were "tools."

Within the laboratories at Western, the flow of mice and mouse resources resembles

the "moral economy" of the *Drosophila* exchange network described by Kohler (1994), where community norms about sharing resources promoted and regulated the transfer of flies between laboratories. Inbred strains of mice are not patented, and mouse strains like the commonly used C57 mouse can be purchased from a commercial supplier and bred without any special permissions. When extra mice are available or another laboratory requests a breeding pair of mice for their own studies, Western researchers typically give mice away for free or for minimal costs that reflect the cost of housing and feeding. They also request breeding pairs of mice from other laboratories in turn, usually offering co-authorship on a paper in exchange for the use of more valuable mouse models such as knockout mice.

The contested commercial status of the mouse in the laboratory demonstrates how researchers' practices can be shaped by forces and trends acting at a distance from the laboratory as well as the interactions of local actors. Responses to events like the patenting of the mouse cancer model known as the Oncomouse demonstrate how the close-knit mouse community has acted to resist and reshape commercial logics of ownership and profit, giving patents new meanings within existing networks of academic exchange (Murray, forthcoming). In response to what they view as a deleterious trend of commodification in academic research, researchers at Western explicitly refuse to make a commercial point of sale part of Mouse X's career, and instead attempt to facilitate his transition between caretakers in a way that resembles a trade or a gift rather than a commercial interaction. Even so, patents and other commercial practices remain part of the landscape of animal behavior genetics. JAX Mice™, for example, are produced and sold by a non-profit corporation, but the trademarked mice are using patented procedures that ensure their genetic stability that in turn increases their value in the mouse community.[3]

[3] For example, the Jackson Laboratory was granted a patent for "Methods for Maintaining Genetic Stability of Inbred Animal Strains" in September 2009 (Jackson Laboratory, 2009).

Day 48: Mouse X and his siblings are sleeping once again when Grace comes into the colony room one afternoon later that week to prepare her newly acquired mice for the upcoming experiment. With the list of cage numbers from Rachel in hand, she pulls cages off of the racks in the colony room and takes them to a workstation. She opens the first cage and "scruffs" a mouse, grabbing him by the loose skin around the back of his neck with her thumb and index finger and turning him over so that he is facing belly up in the palm of her hand. With her free hand, she picks up an instrument that looks like a pair of tweezers and quickly punches a hole in his ear, then turns him over and places him on the wire cage lid where he instinctively grabs for the metal bars. While the mouse is busy scrambling for a good grip on the cage, Grace holds him by the tip of the tail and uses a marker to draw several bands around the base of his tail. When she is finished, she lifts him by the tail and sets him back down on the cage floor. When she reaches Mouse X's cage she grabs him first, and decides to label him number 3. After she is finished, Mouse X has a small hole at the base of his ear and three tail stripes in red Sharpie to distinguish him from his cage mates, and he will now be known as "Mouse 5.3."

Marking mice so that they can be tracked as individuals through the spaces of the laboratory is another step in the mouse's career that prepares him for the transition from mouse stock to experimental subject. Up to this point, Mouse X has a few identifying markings, but not enough for researchers to distinguish him from his cage mates. His brown coat distinguishes him from other strains of mice like C57s that have black fur, his external genitalia have been used to group him into a cage of exclusively male mice, and the card on the front of his cage lists his genotype and his litter's birth date. But in many ways, up until this point in his career he has been treated by researchers and technicians as one animal in what one Smith laboratory technician referred to as an undifferentiated "chunk" of animals, one of a hundred heterogeneous male mice born in July. By marking a number on Mouse X's body, he is now differentiated as an individual for the first time in

his laboratory career. Lynch (1988) argues that indexing animals with a number also helps mediate the animal's transition from a "naturalistic" creature into the mathematical space that the "analytic" animal inhabits.

The work of making mice traceable requires the coordination of different kinds of instruments, practices, and material manipulations to the body of the mouse. With thousands of mice in the laboratory, many of whom look nearly identical, tracking mice is a formidable undertaking that involves multiple systems of digital identifiers, paper records, and methods for physically marking mice. The colony management software that the Smith laboratory uses automatically generates a unique number for each pup born, but Olivia, the information technology manager, explains that there are there are few (if any) links between this digital identifier, the practices of the technicians and researchers, and the physical mouse. When mice are allocated for a particular experiment, researchers typically assign new numbers and whatever ID the mouse had before is lost. Mouse X's new designation—Mouse 5.3—indicates that he is the third mouse in the fifth cage of mice for Grace's experiment. The physical limitations of the mouse, however, make it difficult to make this designation stick, literally. Researchers use tail marking and ear punching in combination because both are highly imperfect systems: Ear punches can become difficult to read after a few weeks and tail markings can rub off even after a few days. Olivia told me that her ideal system for mouse tracking would involve some method for stamping a number impermeably on a mouse so that it would be visible and would never accidentally rub off or be lost. The Smith laboratory even experimented with tattooing identifying numbers on mice's tails, but this method was both too time consuming and limited by the small space available for placing tattoos on the tail.

Day 56: Just a few days before her experiment is scheduled to start, Grace has run into a problem and is meeting with Dennis to discuss her options. She intended to test another strain of mice from the Smith laboratory's colonies at the same time as the mice

that she acquired from Rachel, but the mice are located in two different facilities on campus. While Mouse 5.3 is living in a colony room on the same floor as the Smith laboratory's testing facilities, the second group of mice are in a building that belongs to the associated teaching hospital. Because the facilities are regulated by two different institutional structures, moving the mice from one building to the other would involve a 60 day period of quarantine before the mice could be tested in the Smith laboratory's facilities. If she transfers the mice to the Smith laboratory building, they will be nearly out of the target age range for her study by the time they are released from quarantine. But if she leaves the mice in the other building, she will have to have to borrow testing apparatus from someone because equipment can't be transferred between buildings either.

Leaving the colony room is a critical step in Mouse 5.3's career, because once he is rolled through the colony room door he will be extensively managed as a *disease vector.* At Western, each of the rooms in the facility is assigned a color and a number based on its level of cleanliness. Colony rooms are "green" or "yellow," indicating a high level of contamination control, while procedure rooms are "red," that is at a lower level of control. These designations regulate the flow of materials, mice, and people, directing them in a way that minimizes chances for disease transfer. The general rule is that people and mice can move down to rooms with a lower designation, but not back up. The consequences of disease management protocols are especially important for the flow of mice within and between laboratories and the order of experiments and the work day. While diseases like Mouse Hepatitis Virus typically don't produce serious symptoms in mice, they do severely restrict a mouse's ability to be transported between experimental spaces since other labs are reluctant to accept disease-mice. The disease status of spaces within the laboratory can also have important consequences for the way that researchers conduct their work. At the Jackson Laboratory, I visited a well-equipped phenotyping center that the facility director complained was underutilized because its designation was too "dirty" and researchers

couldn't transport mice back to their laboratories after using the facility. Technicians and grad students at Western also schedule their work days around the cleanliness designations of the various rooms they need to access, working in colony rooms in the morning and conducting experiments in the "dirty" procedure rooms after these tasks are complete.

Mouse 5.3's health status is already being monitored by the veterinarians who inspect him on their rounds and test for some of the more common infectious mouse diseases; but while he remains in the protective space of the colony room it is the staff entering and exiting the room who bear the majority of the responsibility for keeping the mice and themselves disease free. A poster outside of the colony room door with the caption "this is how you should look" shows a properly attired laboratory worker wearing scrubs, gloves, a face mask, and a hair net. When workers leave the colony room, they change out of their scrubs or throw the disposable gowns away before moving into the hallway or a new room to prevent the spread of diseases. But on leaving the colony room, Mouse 5.3's movement will be subject to many more restrictions. He is living in a the room designated "green two" that Sharon jokingly refers as one of the Smith lab's "mouse vaults," and after participating in an experiment in a yellow room he will not be allowed back into the cleaner "green" level room.

Policies that manage the mouse as a disease vector are rarely compromised at Western, and researchers often spoke explicitly about trading off experimental goals against the "practical" problems created by institutional health regulations. For Grace's experiment, compromising the health status of the mouse population at Western by ignoring quarantine rules is not an option. Instead, she must decide which is the lesser of two evils from the perspective of managing her experiments: Testing mice when they are older than her protocol specifies, or borrowing equipment to test them in a different facility at the associated teaching hospital. Both of these options are far from ideal for creating a good experimental mouse, but they are a necessary consequence of the aim of creating a mouse

whose disease status can be monitored and maintained. Grace ultimately decides that testing the mice at a consistent age is more important than testing them in the same facility, and so Mouse 5.3's experiment will take place in the Smith laboratory facilities while his fellow mice drink in another building.

Up to this point in Mouse 5.3's life, he has had little (if any) contact with Dennis or other senior members of the research team. Dennis helped design the experiment that Mouse 5.3 will be participating in, but he has not (and likely will not) physically interacted with the mouse. The work of managing Mouse 5.3 as an experimental subject is "distributed," as Mol (2002) puts it, amongst two different workforces: The senior researchers who design experiments, acquire funding, and write articles; and the graduate students and technicians who conduct the experiments. These divisions of labor, which are common in many behavior genetics laboratories, map onto divisions of space in the physical arrangement of the lab. On the fifth floor of the building where the Smith laboratory is housed, Dennis's office is in a hallway that is easily accessible from the elevator while the colony rooms, wet lab, and procedure rooms are separated by a security door that requires a keycard for entry. The Smith laboratory layout is somewhat unique at Western because their offices and wet lab spaces are in close proximity to the mouse colony rooms, and the meeting room where the Smith lab researchers eat their lunches is directly across the hallway from the animal facilities. This arrangement allows for quick access to the mouse rooms, but has some undesirable side effects since the entire floor smells distinctively (and somewhat unpleasantly) of mouse, even with cage ventilation systems in place. While mice are only a whiff away in the Smith laboratory, mouse housing is ordinarily physically separated from procedure rooms, wet lab spaces, and especially professors' offices, located on other floors or even in separate buildings. Mol argues that divisions of space and labor such as these can diffuse tensions around how to best manage the disease. In the Smith laboratory, these divisions mean that the senior research staff don't have to be concerned with the

management of different kinds of mice in the same way as their technicians and grad students. Laura, for example, had a pet gerbil at home, even though university guidelines recommended not having pet rodents that could harbor diseases that could be spread to research mice. When I asked her about this, she replied that it was not a problem for her since she didn't spend any time in the procedure rooms and therefore never came into contact with the mice. Unlike her technicians, Laura does not have to be concerned with the managing of the mouse as disease vector in this way since she is physically separated from the mice. In some cases, maintaining distance from the physical mouse is more than a matter of convenience or expediency. For several members of the laboratory (including Dennis himself), breaking physical contact with the mouse as disease vector was essential, because they had developed severe allergies to mice over their years of working in the lab. This example also shows how it is not only the mice who are materially altered by the knowledge production process, but the researchers as well.

4.4 The Mouse as Experimental Subject and Ethical Subject

Day 61: Mouse 5.3 is scurrying around his new cage on the first day of Grace's experiment. In preparation for the study, Grace has prepared new cages for her mice and they are now sitting on a cart in the colony room, waiting to be wheeled out to a smaller procedure room. By now, the types of concerns and strategies that she has for managing the *experimental mouse* should be quite familiar. Grace has selected her mice so that they will all begin the experiment at the same age (plus or minus three days), and so that there are an equal number of males and females. She has also divided Mouse 5.3 and his litter mates equally between the experimental and the control groups so that neither group has too many members from a single family. As she wheels Mouse 5.3 and the others down the hallway from the colony room to their new home, she is careful to avoid jostling the cart or exposing the mice to light and noises that might increase their stress levels. As she

arranges the cages in the new room, she randomly distributes their position on the rack to control for differences in light and noise. Mouse 5.3 ends up near the bottom of the rack while his siblings are close to the ceiling. After giving the mice a few hours to acclimate to their new surroundings, she comes back with buckets of water bottles and alcohol-filled sipper tubes. She slots an alcohol bottle in the right side and a water bottle in the left side of each cage, and records the amount of fluid in each bottle. In the coming days, she will return to the room to read the levels of the bottles and switch their positions so that the mice don't develop a "preference" for drinking from the left or right side of the cage.

Each stage of Mouse 5.3's life over the next six weeks has been carefully laid out in both Grace's experimental protocol and in another document called the Animal Component of Research Plan (ACORP). In addition to being an experimental subject, Mouse 5.3 now also has the status of an *ethical subject*, whose future pain and suffering has been weighed against the data that this experiment is expected to produce. Mouse 5.3's transition to an experimental subject is clearly indicated through procedures like tail striping and "assigning" him to a researcher, but how Mouse 5.3 acquires the status of an ethical subject is less clear. In one sense, Mouse 5.3 has always been marked out as a different kind of mouse. As Herzog (1988) points out, there are often several different kinds of mice in a laboratory setting, each with a very different ethical status. In his psychology laboratory, most of the mice housed there are treated as experimental subjects whose lives and deaths are regulated by the IACUC, but there are also mice kept as live food for other experimental animal subjects and pest mice in the hallways who are killed with mouse traps of questionable efficacy and that institutional review boards would never permit (Herzog, 1988). Although Mouse 5.3's career advancement to the status of an experimental subject was never certain, his status as a *potential* experimental subject means that he always also been ethically differentiated from other animals.

The ACORP that describes the experiment that Mouse 5.3 has been enrolled in has also

already been reviewed by the Institutional Animal Care and Use Committee (IACUC), a committee of scientists, veterinarians, and laypeople that review all protocols involving animal subjects at Western. In the documents that they submit to the IACUC, researchers describe how many mice they plan to use, why they are using mice and not another species, what procedures they plan to do with the mice, and whether these procedures are likely to cause the mice discomfort or pain. When the protocol is reviewed, members of the IACUC can comment and make recommendations about how animals should be cared for (for example, what kinds of drugs might be given to mice after a surgery and how often they should be checked) as well as judge whether there is an adequate scientific justification for the number of mice and the kinds of experiments that researchers are proposing. The committee also makes recommendations about the types of procedures that researchers should use for their experiments, such as how the mice should be euthanized at the end of the experiment. These recommendations may differ between different IACUCs, and so while Mouse 5.3 will have his blood sampled through a small nick in his tail, his fellow experimental subjects in the university hospital building will have their blood samples drawn directly from the vein on the recommendation of the hospital IACUC. If the protocol is approved at the end of the meeting, the principal investigator is granted approval to enroll a specified number of future mice in the study.

At the institutional level, much has changed in the way that laboratory animal research is regulated in the United States in the past four decades. Many of the bureaucratic requirements and institutional structures like IACUCs and ACORPs that regulate animal welfare today didn't exist when senior researchers like Dennis started their professional careers. The Laboratory Animal Welfare Act, first passed in 1966, charged the United States Department of Agriculture with the oversight of animals used in laboratory research, but many commonly used research animals (including mice and rats) were excluded from these oversight procedures. In the 1970s, NIH policies required researchers using federal

funds for research with live animals to show that they were accredited by an animal welfare agency or had their own committee to oversee animal research (including research with mice), and in the late 1980s the Animal Welfare Act was amended to include a provision mandating the establishment of institutional review committees to oversee animal research, modeled after the Institutional Review Boards for monitoring human subjects research (Rozmiarek, 2007). These institutional bodies, which came to be known as the Institutional Animal Care and Use Committees (IACUCs), monitor research with many different species and ensure that the researcher is in compliance with other regulations like NIH policies as well as the Animal Welfare Act. The existing laws and regulations are summarized and interpreted in the *Guide for the Care and Use of Laboratory Animals* (known in the animal welfare world simply as "the Guide"), a publication jointly authored by institutions such as the National Research Council and the Institute of Laboratory Animal Research to help researchers, veterinarians, and IACUCs make recommendations for improving animal welfare. In addition to institutional review committees, both public and private institutions can undergo voluntary accreditation by agencies like the non-profit Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC), founded in 1965, to show that they adhere to the standards outlined in the *Guide*.

This changing regulatory landscape suggests that there is potential for conflict between these new institutional structures and existing laboratory practices, but researchers at Western often de-emphasized any negative impacts of institutional animal welfare requirements on their practice (other than the unwelcome chore of having to do more accounting and paperwork). Aiden, a member of the animal care staff, explains that this is exactly how the animal welfare oversight mechanisms are ideally supposed to operate:

> *Aiden:* The laws and regulations are summarized and interpreted for the IACUC by the *Guide*, and the *Guide* establishes them for the institution and essentially acts as a shield. It shields the people working from the large array of things that they could worry about, and also it shields the people worrying about regulation from the people that are actually doing the work. It acts as a

stopgap there.

NN: So if somebody's going to take flack for a protocol that was approved by the IACUC, it's going to be the IACUC and not be the investigator.

Aiden: Absolutely right.

In Aiden's description, the *Guide* mediates between the complicated landscape of animal welfare laws, regulations, recommendations and those who are responsible for enforcing those laws and regulations on the IACUC. Likewise, the IACUC and its approval procedures provide an organizational separation, dividing labor of animal welfare from labor of performing experiments. These separations ensure that the researchers themselves are not held directly responsible for knowing which regulations apply to their work, and they are also intended to make day-to-day life easier for researchers who work with animals. Aiden explains:

> What [the IACUC] does is it provides a mechanism for people doing the work every day to not concern themselves with the laws and regs. So the PI submits a protocol, the protocol is accepted by the IACUC. Lots of things happen in between there as you saw, but when those two basic things happen, all of a sudden all that the person doing the work has to do is open, read the page about what they're supposed to do, and do that. If they do that, they're complying with all the existing laws and regulations.

Aiden portrays the ideal functioning of the IACUC system as a kind of ethical "black box" where researchers working in the laboratory can use the decisions of the IACUC as a means of ensuring that they are meeting the standards for ethical treatment of their animals without becoming experts themselves on animal welfare legislation. As long as they are performing procedures approved by the IACUC, they will be acting within the bounds of existing regulation and legislation about the ethical use of animals in research.

The kind of separation between the role of the experimenter and the role of the IACUC members that Aiden describes seems somewhat implausible since scientists from Western form a substantial part of the committee membership, and yet scientists themselves also talk about their participation in these activities as separate from their experimental work.

Elizabeth, who was the chair of the IACUC while I did fieldwork at Western, is also a principal investigator in the department of neuroscience. When I approached her about observing some of the meetings, she asked me why I would be interested in sitting in on the IACUC meetings. I told her that I was studying methodology in animal behavior genetics, and I was interested in seeing how IACUC proceedings played a role in forming research methodology. She looked at me with a puzzled expression and told me that the IACUC didn't deal with methodology, it dealt with animal distress. I was quite surprised by this response: To me the IACUC had a clear role to play in shaping how new methods developed since it determined which kinds of techniques and procedures could be used on mice, even if those methodological choices were based on animal welfare concerns and not scientific concerns. To her, calling these kinds of decisions "methodological" decisions was equally strange.

Elizabeth's reaction suggests that it is possible for ethical and experimental mice to be "co-produced" at Western in ways that researchers experience as largely unproblematic, and even mutually reinforcing. The IACUC generates lists of ethically justifiable procedures, and scientists building research protocols within those constraints. In the contested terrain of animal research, doing animal behavior genetics research also requires the co-production of structures for evaluating and governing the use of animals. Without these structures, experimental work could not take place. Researchers would not be able to secure federal funding, use institutional facilities for housing animals, or counter the concerns of protestors who allege that experimental animals suffer unnecessarily. Objections to the use of animals in research still exist, but the social practices for governing research stabilize social concerns about some kinds of laboratory practices so that research can take place.

The perceived separation between the role of the IACUC and experimental practice does not mean, however, that the Western laboratory is devoid of discussions about the

ethics of animal research—quite the opposite. All researchers who work with animals at Western and other universities also take some training courses that familiarize them with existing welfare regulations and the rationales behind performing certain procedures and not others. Western's "Working with Mice" training module explains the rationales behind many existing regulations (such as why the IACUC recommends that researchers use solid floors instead of wire mesh for mouse cages), and also educates researchers about how to make their own animal welfare assessments (such as ways to tell if a mouse is still in pain or distress after a procedure).

Developing a position on animal research ethics is also an important part of the socialization process of becoming a successful practitioner in the sometimes controversial field of animal research. In addition to co-producing institutional structures for governing the ethics of research and experimental practices, ethical researchers and experimental scientists are also being co-produced at Western. In interviews about the training process at Western, many students told me that one of the most valuable things that they learned in their rotations through the primate center in their first year was how to talk about the ethics of animal research. Senior and junior researchers openly discuss the problems that they face when trying to talk about the fact that "mommy kills mice for a living" to their friends and family members, as one researcher put it. Finally, researchers also described a process of coming to terms with the ethics of animal research for themselves. Emily, who did research in a laboratory that studied human subjects before starting in the graduate program at Western, recalls that dealing with her own emotional reactions to doing experiments with animals was a difficult part of her first year at Western. She says:

> I mean, I had a hard time adjusting when I first got here, because I've never worked with animals, and it's not easy to do some of the things that we do. And I think if it were then we shouldn't be, my opinion is that if it's too easy to do some of these things, then you shouldn't be an animal researcher, because you're not paying enough respect to the animals. There have definitely been very hard times, but in the end you just have to believe that what you're doing is going to make people's lives better.

Emily describes her feelings as a normal response to her first experiences doing research with animals, and even suggests that scientists who do not experience such reactions are not well equipped to become animal researchers.

In his study of ethical practices in the more contentious field of nuclear weapons development, Gusterson (1996) observed that weapons scientists also talked about the process of coming to terms with the ethics of their research in individualistic terms. But he argues that despite the fact that researchers talk about their opinions as highly personal and something that they have developed their own, they all conform to what he calls the "central axiom" that developing better weapons helps stabilize nuclear deterrence (p. 56). Researchers at Western are similarly socialized into a world where the decision to use animals for research is often portrayed as something intensely personal that each individual needs to think about as an individual. Scientists had to decide for themselves if they were willing to perform certain procedures, or even work with animals at all. Liam, for example, said that "trying to justify why I do that is something that I do struggle with on my own." Researchers describe how working with animals involves personal decisions that they had to make about their capacity for doing research with animals, and that different researchers might make different personal assessments. And yet, most researchers' positions hold up what might be called the "central axiom" of animal research: the utilitarian argument that the harm to animals is outweighed by the greater good of combating human diseases. Ava's comments nicely illustrate the two facets of scientists' ethical self-expression that Gusterson (1996) describes:

> Like me personally, I could never work with monkeys. I would have a really hard time doing that. But I don't think that monkey research is bullshit and it shouldn't be happening. Quite the contrary. I just know that I absolutely should not be doing it.

Ava describes her decision not to work with monkeys as as a personal decision, while at the same time affirming that research with monkeys is a valuable and needed component of animal behavior genetics research.

These ethical selves and experimental selves that researchers at Western develop are sometimes both at work on the mouse in the laboratory. Alex, for example, describes his concerns about his first time conducting a new set of experiments in both ethical and scientific terms. He has just started a series of experiments that involve injecting a receptor-blocking drug directly into a rat's brain, and is having a difficult time dealing with his emotional response to the pain that these injections cause his rats:

> *Alex:* Fast forward to the day that you give an injection, a microinjection, and apparently there are people out there that have gotten so used to it that it's no big deal. Or maybe they just habituate the animals to these injections better than I was doing it by repeated probing or whatever. But the rats that I was microinjecting in the series of experiments that I've done so far, they just, I just couldn't take it. I was almost in tears over these guys, because they hate it.
>
> *NN:* You can tell that they're in pain or they don't like it or whatever?
>
> *Alex:* Yes. They're squirming, they're screeching … And you talk to people and they'll tell you, oh you know, what you should do is wrap a towel around them so that you get a better grip and they don't fight as much. And it's like well, it's not the fighting that's the problem with me, it's the fact that I know I'm causing these animals physical pain right now, and that just really bothers me. And it doesn't just bother me from a moral perspective … I don't know if this is what you had in mind, but it bothers me from an "oh my God, I'm just creating so much stress in these animals right now and then I'm going to go and test them!" And I'm going to … how can I, you know, realistically say that that super stress that I just gave it, a two minute injection where they're squirming and squealing the whole time, and then I put them in the box and say "hey, show me what you've learned but don't let stress affect you." It's just ridiculous.

Alex distinguishes the ethical concerns that he has for the subjective experience of the rats he is working with from the experimental concerns he has about how this stress will affect his experiment, but he acknowledges that he is bothered by both. He is finding that the practices he has available to him are not up to the job of managing either his ethical concerns about the experiment or his experimental ones. Wrapping the rat in a towel may make it easier to get the rat to comply as an experimental subject, but it doesn't do anything to enhance the welfare of the rat. He mentions that habituating the animals

to the injections may be one way to simultaneously enact an ethical and experimental animal, but he is unsure whether this will fully address his concerns about the welfare of his animals.

4.5 Environmental Enrichment and the Career of the Mouse

Day 70: For Grace, managing her ethical concerns about Mouse 5.3's participation in her drinking experiment is a little easier. The experiment has been running for a little over a week, and today I accompany her as she makes her rounds through the experiment room to check on her mice and refresh their alcohol bottles. Grace wheels a cart with fresh bottles of alcohol solution into the room, and as she places a bottle on each mouse cage she calls out the volume of each bottle out to me to record on a chart on a clipboard. Now that he is participating in an experiment, Mouse 5.3's accomodations are somewhat different from his former cage in the colony room. Instead of being housed with littermates, he has been moved to a separate cage so that Grace can measure how much he drinks as an individual. As we pass through the rows of mice I notice that his cage also contains a square of cotton fluff that is often given to breeding mice to build nests. I ask Grace about this "nestlet" square, and she tells me that all of the mice in the room have nestlet squares because they are individually housed for her study, and the bedding material reduces the stress of being housed alone. She suggests that giving nestlet squares to singly-housed mice may be an institutional policy, and points out that even the "sentinel" mouse, a mouse that is kept in a separate cage and tested regularly for signs of disease, has a nestlet square provided by the animal care staff.

The nestlet square that has been added to Mouse 5.3's square is what is known in the laboratory as an "enrichment device," defined as any object that is added to a mouse cage other than food, water, and bedding. Some of the enrichment devices that are commonly used at Western include nestlets and plastic shelters shaped like a toilet paper roll tube.

Other more elaborate enrichment practices include giving mice larger cages (often rat-sized cages), adding sticks and twigs to the cages to recreate more "natural" environments, providing exercise equipment such as wheels, introducing new objects and removing old objects on a regular basis to create a continually changing landscape within the cage, or providing connecting tubes between cages so that mice can move freely through a system of cages (Marashi, Barnekow, Ossendorf, & Sachser, 2003).

The increasing use of enrichment objects in rodent laboratories is one instance where some researchers describe the interaction between welfare practices and experimental practice as potentially disruptive. The effects of introducing "enrichments" to the housing environment has been a topic of interest in rodent research for many years, beginning in the 1950's with psychologist Donald Hebb's observation that the rats he gave to his children as pets performed better on learning and memory tests than the rats he kept in his laboratory (Hebb 1949, cited in T. Schneider, Turczak, & Przewłocki, 2006; Würbel & Gardiner, 2007). Psychologists such as Mark Rosenzweig continued this line of research, looking at how environmental enrichments or impoverishments impacted the development of the brain and the performance of animals on learning and memory tests (see, for example, Bennett, Diamond, Krech, & Rosenzweig, 1964). Studies on environmental enrichment also provided some of the first opportunities for exploring gene-environment interaction. One research team, for example, showed that performance of "maze bright" and "maze dull" rats bred for their learning abilities could be altered by raising them in enriched or impoverished environments (Cooper & Zubek, 1958). Beginning in the 1990s, there was a resurgence of interest in environmental enrichment for laboratory mice and rats (Hutchinson, Avery, & Vandewoude, 2005). Researchers working on neurodegenerative diseases like Alzheimer's disease and Huntington's disease have been especially interested in exploring environmental enrichment as a way to enhance the plasticity of the mouse (and potentially human) brain; and there is strong enthusiasm in these fields for environ-

mental enrichment as a model for treating neurodegenerative diseases, especially because few other treatments are currently available (Spires & Hannan, 2005).

While most of the literature on environmental enrichment for rodents focuses on experimentally altering neurological function, there is also a growing body of literature on how these same techniques can be used to improve animal welfare. A review of the enrichment literature shows that approximately 20 % of the articles published on cage enrichment since 2001 focus specifically on animal well-being (Hutchinson et al., 2005). Environmental enrichment aimed at improving animal welfare is not yet explicitly part of the IACUC's welfare recommendations at Western, but it already happens in ad hoc ways in Western laboratories. In the Smith laboratory, researchers use objects like mouse huts and nestlets for environmental enrichment experiments, but they also use them on some occasions for mice that are difficult to breed or mice like Mouse 5.3 that are housed alone. A recent survey at the National Institutes of Health showed that environmental enrichment is widely but unevenly practiced in many kinds of laboratories that use mice. All of the animal facilities surveyed at the NIH provided enrichment to at least some of their mice, and almost two thirds provided enrichment to all of their mice (Hutchinson et al., 2005).

One of the reasons that tools and findings are being repurposed to achieve welfare aims by the animal welfare practitioners is that the welfare community lacks a foundation of tests and findings of its own. Amy, a veterinarian who is responsible for developing and implementing animal welfare policies at a major commercial mouse supplier, explains some of the problems that she encounters in her work:

> You know, one of the quandaries that we have from a regulatory standpoint is that in the Animal Welfare Act there's a specific contingency for canine exercise programs. So if you have dogs in your laboratory, you have to have a documented exercise program for them. However, no one has ever shown how much exercise, how often, what kind is beneficial for the animals. You know, is it better if we give them just really big runs so they can run around whenever they want? Or do they really need time out of the enclosure? …

> The idea is right, but no one to my knowledge has quantified what is enough, what is not enough, what is beneficial.

For many aspects of animal welfare, little experimental data exists for welfare practitioners to draw on when making recommendations. Tests developed by animal behavior geneticists to use mice to model the effects of stress in humans may not be a perfect fit for welfare practitioners, but they are often the closest available thing to a test to measure "stress" in mice.

For animal behavior geneticists, the overlap between these practices creates confusion about what kind of ideal mouse is being enacted when researchers put a plastic mouse shelter in a cage in their laboratories. Ethan, a graduate student at an East Coast university, says that he thinks the idea of adding enrichment objects into mouse cages makes a lot of sense, since laboratory animals are taken out of their natural habitats and housed in highly artificial conditions:

> The idea that you're taking mice that sort of evolved in this certain ecological niche, and you're putting them in these little cages, housing them in very sort of distinct social groups, that's in a lot ways a very unnatural environment for them. And of course that's going to effect their behavior among other things. And I guess the question I struggle with is what does that matter, right? Like what is the... who's to say that a mouse reared in sort of wild, natural environment is more or less apt for modeling human conditions than a mouse reared in a cage? And that's … I don't know. On some really just sort of a hunch level I feel like my experiments might be better if I was using, you know, a big group housing environment somehow. But then on the flip side you're also increasing environmental noise.

Ethan is interested in the idea of environmental enrichment, but he is uncertain about what the widespread use of enrichment practices in mouse experiments would mean for the status of the experimental mouse. It is not clear to him that a mouse housed in an enriched environment would be a more ideal experimental subject than a mouse housed in a standard cage. He explains that he would like to try some environmental enrichment techniques in his own practice, but thinks his efforts are likely to be misinterpreted by the scientific community:

> When I submit that article, people are going to say, "oh, you did an environmental enrichment experiment," you know what I mean? Because it's not sort of standard rearing. So you know that's something that immediately would stand out about that paper, and people would say … it would become a point of discussion. Whereas on some level I feel like I could say, well … "you're doing experiments in environmental deprivation. I'm just doing a normal experiment."

Ethan suspects that his colleagues' first reaction to his use of enrichment practices will be to assume that he is using them as experimental manipulations, especially if he reports on them in a paper.

While Ethan is cautiously optimistic about the adoption of enrichment practices in normal scientific work, not all researchers at Western are quite so enthusiastic. Dennis sees the use of the environmental enrichment research to support welfare programs as a misappropriation of the scientific literature:

> There are people running around who assert that the standard laboratory environment amounts to cruel and unusual punishment for mice, and that therefore you have a moral obligation to give them an enriched environment by piling in toys until there's no bedding left, and that's going to make them better and happier and less weird. There's remarkably little data to back that up. There's quite good data on environmental enrichment and learning in rats in a very specific sense, but not in the kind of general animal welfare sense.

Dennis objects to other practitioners' uses of environmental enrichment techniques and findings for purposes other than what he thinks that they were designed to do. For him this amounts to a mis-appropriation of the epistemic scaffolds of a particular test to support claims that it was not designed to carry. In his view, the environmental enrichment experiments conducted in the 1950s and 1960s were conducted with a specific purpose in mind: to demonstrate the effects of the environment on learning and memory, and not to give information on the subjective well-being of laboratory rodents.

Linking environmental enrichment to animal welfare also could have future consequences for the way that mouse welfare is regulated in the laboratory. The rhetorical flip side of "enriched environments"—"impoverished environments"—makes it difficult for

researchers to argue against the widespread introduction of enrichment devices in rodent laboratories. Hanno Wï¿½rbel, an ethologist at a university in Germany, makes a forceful case for the increasing use of environmental enrichment practices by arguing that these practices are not only good for science, but that *not* using these practices could jeopardize both the ethical care of the mouse and the validity of scientific research. He argues that the standard cages that mice are typically housed in have adverse effects on their behavior, making them poor models for humans:

> My research has shown that these animals go nuts, basically. They develop abnormal behaviors, and there's a number of signs that their well-being is impaired, seriously impaired, not just they're not jumping or running in circles or gnawing the bars because they're bored, but it seems to reflect … really mental disorders in these animals. And why these animals should be better models for human studies, animals that suffer from severe behavioral disorders, is a total mystery to me.

Veterinarians and animal welfare advocates may appreciate Wï¿½rbel's message that welfare and science go hand in hand, but not all researchers are so enthusiastic about the transition that Wuerbel is fostering between ethical and experimental practices. While we often think of "boundary objects" as devices for stabilizing the connection between two social worlds, connections between social worlds might also be unintentional or unwelcome (Star & Griesemer, 1989).

4.6 The End of the Career of the Laboratory Mouse

Day 103: Six weeks have passed since the beginning of Grace's experiment, and while Mouse 5.3 licks his bottle of alcohol alone in his cage, Grace is meeting with Sharon to decide his fate. Grace's experimental protocol has been completed, but she wants to keep collecting data for a few more weeks. She is thinking of continuing the protocol to see if Mouse 5.3's drinking patterns change as he gets older, or she might try altering the concentration of alcohol that she gives him in his daily bottle. Alternatively, she suggests that she could use Mouse 5.3 and his cohort in a pilot study to test a new way of attaching

the bottles to the cages, to see if this would help eliminate the side preference that the mice developed during her experiment. Sharon finds these requests a little surprising. The protocol calls for the mice to be euthanized at the end of the experiment, and she wonders why Grace wants to keep using these mice instead of starting a new experiment with a fresh batch of mice.

Grace and Sharon's deliberations around how to end the career of the laboratory mouse demonstrates how institutional contexts can subtly shape researchers' conceptions of acceptable experimental practice and the ideal experimental animal. When I ask Grace about why she thinks Sharon found her request odd, Grace attributes their difference of opinion to differences in the financial resources that Western researchers have available compared to her home university in Europe. She explains:

> For example, I'm supposed to work with mice who are not too expensive, so for example C57s are very expensive, so if I design an experiment it's better if I can use another strain. And if we don't have enough money, it's good to do an experiment according to the cheaper strain, you know what I mean? So I don't choose sometimes the strain according to my objective, I change my objective according to the availability of the strain.

Grace, who is used to operating in an environment where funding for experiments is limited, has developed customs for selecting and using mice that are different from that of the researchers at Western. For her, the cost of the mice was highly salient, while the different distribution of resources at Western made space highly relevant and the commercial status of the mouse almost a non-issue. But even though she suggests that she designs and executes her experiments differently in a comparatively less well funded laboratory, she argues that the conceptions of what counts as good science in both the Smith laboratory and her home laboratory are the same. She says:

> Grace: But I'm surprised to see that in [European country] or here, except for the money, the way we think is exactly the same. I mean, the same idea, the same … yeah. Sometimes I thought that because I was in a small country with not so much money, we made cheaper things, but no. Not at all. We make cheaper experiments, but not cheap science, do you know what I mean?

> NN: Right, right. So it's the same kind of thinking.
>
> Grace: Exactly. It's … very serious in [European country], like here. When I do an experiment in Europe, it's the same to the experiment that I do here.

But as we keep talking, subtle differences emerge between her way of conceptualizing an acceptable experimental mouse and the way that Sharon and other researchers at Western conceptualize acceptable experimental practice. Grace explains her rationale for why she wants to use the mice in her experiment for longer than the protocol dictates:

> Because of that, if I use mice I'm supposed to use them for one experiment, of course, but after, if I have something, I could use the mice a second time, just to test a new idea, you know? … You are supposed to use them until the end, so if for example I do an experiment with seven injections of whatever, after seven days I have my data, but just for fun I do to maybe 20 injections, just to be sure that I have all the data possible from these mice. So I use them until the end, because it's expensive. When I say, okay I will perform my experiment and then maybe I could keep my mice for a few more weeks just to see what happens if I do this or this, just for fun, they look at me like this [makes a face]. Because nobody does that here. So if I want to test another hypothesis, I have to take another bunch of mice and do another experiment.

Since mice are an expensive resource, she was trained to use mice "until the end" and extract as much data from them as possible. But the idea that mice should or could be used for as many experiments as possible also conflicts with ideas at Western about what constitutes good experimental practice. Mouse 5.3's prior experience in Grace's drinking study makes him a less than ideal experimental mouse, and other researchers at Western suggest that Grace should just take a fresh batch of mice rather than generating data whose quality is suspect with Mouse 5.3. Grace's committment to use her mice "until the end" could also be interpreted as something that not only makes good economic sense, but is also respectful to the life of the mouse. The idea that using mice ethically requires using them as long as possible is in contradiction to the IACUC's recommendations that mice should be given a "good death" as soon as the experiment has reached a conclusion. Thus, even though Grace attributes the difference of opinion between her and Sharon to a difference in perception of Mouse 5.3's value as a commercial object, there are also

perceptible differences between conceptions of the ideal experimental and the ideal ethical mouse visible here.

Day 117: Following her conversation with Sharon, Grace continued on her drinking study. But for Mouse 5.3, the reprieve is temporary. Two weeks into her continued experiment, Sharon tells Grace that another researcher needs the procedure room that she is using for an upcoming experiment. While Mouse 5.3's drinking behavior and genetic information have already been immortalized in Grace's experiment, the career of the living, breathing Mouse 5.3 comes to an abrupt end. One hundred and seventeen days after his birth in a colony room down the hall, Grace euthanizes Mouse 5.3 by quickly and skillfully breaking his neck in a cervical dislocation.

4.7 Conclusion

In this chapter, I attempted to take a broader view of how experimental systems in animal behavior genetics are developed and stabilized by returning to the arena of the laboratory. Tracking the trajectory of a single, fictional mouse—Mouse X / Mouse 5.3—as it moves through its "career" in the laboratory traces all of the many actors, regulations, institutional bodies, and discourses that act on the body of the laboratory mouse. Throughout the course of his career in the laboratory, Mouse X has always been managed with his potential future career as an experimental subject in mind. His genetic background and early environment were managed for consistency so that he could be compared to other past and future experimental mice even before he was selected for Grace's study. Once enrolled in an experiment—and becoming Mouse 5.3—his surroundings were intensively managed to control for factors that might change his behavior other than the ones that researchers were interested in testing. But alongside the experimental subject that scientists have been carefully grooming from birth, other actors are also acting on the body of the mouse

and assigning different statuses to it. Armed with ear punch tweezers, computer records, scoops of food, and welfare regulation guides, different members of the laboratory are also working to create other kinds of ideal mice that are not always exactly the same as the ideal experimental mouse.

While it is easy to think of the laboratory mouse as being primarily (or entirely) an experimental subject, I have argued in this chapter that many other ideal mice are being managed alongside the experimental one. At various times in his career in the laboratory, Mouse 5.3 was at once a breeder, a disease vector, a commodity, a traceable entity, an ethical subject, a data point, and of course, an experimental subject. Identifying these different ideal mice and the practices that constitute them is useful because all of these practices have the potential to shape the form that the mouse as experimental subject eventually takes. The behavior genetics laboratory might be thought of as an ecological system where people and mice with different and sometimes competing needs exist in different "niches." In some instances, the multiple ideal mice of the behavior genetics laboratory seem well adapted to their particular niches and co-exist with minimal conflict. Different workforces and spaces distribute tensions between the different actors that interact around breeding mice in the colony rooms, so that researchers (especially senior researchers) rarely talk about the breeding mouse as something that impacts their research. Likewise, the mouse as experimental subject and the mouse that is regulated by the IACUC are co-produced by researchers. Even though in some cases the same individuals are both designing experiments and evaluating similar experiments in the setting of the IACUC, researchers often talk about animal welfare regulations as fitting seamlessly with their experimental work rather than conflicting with it.

In other cases, researchers decribe instances where the different mice in the laboratory compete and conflict with each other. When researchers at Western talk about the "practical limitations" of what they are able to accomplish in the laboratory setting, they are often

referencing conflicts between the ideal experimental mouse and the mouse as disease vector. Western University's policies for disease control and the distributed configuration of the campus buildings and regulatory systems means that the mouse that is managed by these different policies often comes into conflict with the ideal experimental mouse. The competing needs of these two ideal mice are quite visible to researchers at Western, and they acknowledge that sometimes the ideal experimental design might be altered because of these policies. In other laboratories, different kinds of tradeoffs may be more prominent. In a laboratory where researchers are working with mice that are more difficult to breed or where they have more limited funding, conflicts between the experimental mouse and other kinds of mice might appear more often in scientists' concerns.

Finally, looking at the experimental mouse and the ethical mouse in the behavior genetics laboratory in particular shows how shifts in the landscape of the laboratory can disrupt existing understandings of what constitutes an ideal mouse. Environmental enrichment techniques, which were originally developed by behavior geneticists for studying environmental effects on learning and memory, are increasingly being used for other purposes by researchers and veterinarians in mouse laboratories. Using environmental enrichment to improve the welfare of the mouse rather than for scientific study creates ambiguity about what the ideal experimental mouse should look like. While some researchers see the use of enrichment as a way to accomplish experimental and welfare objectives at the same time, others see the potential slippage between welfare and experiment as unwelcome.

5 Communication Complexities: The Public Life of Behavior Genetics Facts

In the final weeks of my semester at Western University, Dennis and I met in his office to discuss the upcoming session on "ethics" in the introductory behavior genetics class. As the resident sociologist, Dennis had asked me to recommend some readings on the topic. He browsed through the list of readings that I had compiled, and stopped at two popular articles from the *New York Times* and *Harper's* magazine. These kinds of popular articles "can really make behavior geneticists crazy," he commented, because he thought they were typically overly optimistic or overly pessimistic about the future of behavior genetics research: either they make grandiose claims about the field's potential to produce new drugs or genetic susceptibility tests, or they argue that it will never be possible to link specific genes to behavioral disorders.

When I asked other researchers at Western about the public communication of their research almost all agreed that it was a necessary and important component of scientific life, and yet many discussed the topic with an obvious sense of unease. Scientists in many fields talk about public communication as something that is not only difficult but "dangerous" (S. R. Davies, 2008). As Dennis put it:

> It's a tough game. You're damned if you do, you're damned if you don't. You can't control the message, you try to control the message, and you know, scientists who deal with the press learn the hard way—or if they're lucky, by having good media people to work with—that you really, really have to be careful in packaging a scientific message for the public, for the press. Because the press will just eat you alive, they've got their own agendas.

Many members of the behavior genetics community share Dennis's feelings on reporters and the popular press. As the authors of one *Science* magazine opinion article put it, "the genetics of behavior offers more opportunity for media sensationalism than any other branch of current science," and they bemoaned the fact that news reporters exacerbate the problem by publishing headlines claiming that researchers have discovered the "gene for" complex traits like homosexuality, aggression, or risk-taking (McGuffin, 2001, p. 1232).

These anxieties are not limited to the human behavior genetics field. Animal behavior geneticists also feel that their subject matter is one that is frequently sensationalized, criticized, misinterpreted, and misrepresented in public opinion and the popular press, making the task of communicating their research to the public difficult. As one practitioner put it to me, what makes talking about animal behavior genetics to non-scientists problematic is the "animal," the "behavior," and the "genetics." Researchers talk about the "long scale war" that they are engaged in with animal rights activists about the utility of animal research, and researchers at Western in particular express reluctance to talk about their work with friends or acquaintances because animal rights groups are very active on the west coast of the United States. Even mouse, rat, and fly researchers who are not typically the targets of protests describe negative reactions to their work from non-scientists, and feel an obligation to publicly defend animal research on behalf of the field as a whole. Researchers are also concerned about how to talk about the heritability of behaviors to a public audience. The social salience of research on behavior and past controversies around the heritability of intelligence make behavior geneticists worry that putting forward results about the genetic basis for certain behaviors in a public setting could go "very wrong, very quickly," as one researcher put it to me. Finally, researchers express concerns about the public's negative associations with any kind of genetic research, which is haunted by past histories of eugenics and Nazi science, and contemporary concerns about privacy and control of genetic information. Emily, who worked in a laboratory using human

DNA samples prior to coming to Western, recalled that study participants were often quite suspicious about the use of their genetic material. She said that participants had the impression that researchers might be "doing something shady with their genes," like using them for cloning or "putting them in some kind of secret government database," despite her assertions that the DNA samples would be destroyed after the study was completed.

Throughout this dissertation I have explored the ways in which researchers craft their claims by calibrating the strength of the relationship that they draw between the animal and the human, and between genes and behavioral traits. Discussions around how to control genomes and environments and how to use animals to model specific aspects of human disorders are all part of ongoing discussions in the scientific community around what researchers can reasonably expect to say about the relationship between genes and behavior, but these questions are also isomorphic to questions raised in public debates about how to use genetic knowledge to understand humans: What kinds of statements can be made about the relationship between genes and behaviors, and what kinds of modifiers and qualifications are needed? If "gene for" formulations will no longer work to describe these relationships, then what will be put in their place? How much predictive value will findings have, and in what ways should behavior genetics facts be used to form public policies or clinical interventions? These kinds of questions about how genes function, how stable behavior genetics facts are, and how quickly knowledge will accumulate are at issue in methodological debates within the scientific community, but they are also questions that are central to public debates about genes and behavior. Shapin and Schaffer's well-known argument that "solutions to the problem of knowledge are solutions to the problem of social order" calls attention to the deep connections that often exist between methodological questions about how to produce acceptable knowledge and political questions about how to order societies. (Shapin & Schaffer, 1985, p. 332). They argue that methods of knowledge production such as "the experiment" are tied to particular visions of the polity

that scientists operate in, and to visions of how the products of these knowledge production systems will be used in the public sphere. Jasanoff (2004) has likewise demonstrated that discussions in different nation states about how to evaluate and govern contentious objects such as genetically modified organisms and stem cells are inflected with different notions of citizenship, deliberation, and accountability.

This final chapter examines how animal behavior geneticists themselves envision and participate in these public discussions about the meaning and implications of behavior genetics facts. In this terrain of contentious debates and tainted past histories, communicating "complex" behavior genetics facts to the public can appear to be a nearly impossible task to behavior geneticists. For practitioners, finding ways to represent the uncertainties that they see in behavior genetics findings and methods is even more difficult in the public sphere than it is in the scientific community. On the one hand, researchers anticipate that strong claims about links between genes and behavior made in public venues could be both scientifically and socially dangerous, negatively impacting the credibility of the field and providing opportunities for undue speculation about of the power of genetics. On the other hand, behavior geneticists fear that if they use too many qualifiers and make too much room for uncertainty in the ways that they present the relationship between genes and behavior, they risk looking "wishy-washy" and unscientific to their scientific colleagues and to the general public.

In the first two sections of this chapter, I examine how behavior geneticists attempt to manage the public image of their field, and how they envision "the public" that responds to behavior genetics facts. I describe how researchers construct accounts of their field for the public by defining behavior genetics against past histories of eugenics research, ideological uses of behavior genetics facts, and visions of the future clinical applications of behavior genetics findings. Researchers do not necessarily disconnect behavior genetics from controversial past histories or future applications entirely. Past histories of eugenics,

for example, may become warnings about how behavior genetics may be misused; or concerns about how behavior genetics may be applied in the clinic may be lessened by pushing expectations for applications into the distant future. Researchers also construct representations of the public that responds to behavior genetic research and the political culture in which they work. I argue that researchers draw on specific and local resources in order to generate these representations of the public, and show some of the ways that researchers at Western draw on past debates about the heritability of intelligence, the animal rights movement, and even public discussions about evolution and global warming in order to anticipate how the public might respond to their work.

Thus, I argue that both the public image of behavior genetics and Western researchers' understandings of the political culture that they work in are structured by contentious and polarized debates, and the resulting model of how science communication works is not a terribly optimistic one. In researchers' estimations, some of the publics that respond to behavior genetics facts are not only likely to be lacking in their knowledge of how genes work to produce behavior, but they are also potentially capable of intentionally misrepresenting or distorting behavior genetics. In the final section, I explore how behavior geneticists envision their roles as managers of the public life of behavior genetics facts. As I alluded to at the end of chapter 2, researchers calibrate the strength of the claims that they make about the relationship between genes and behavior not only for scientific publications but also with other audiences in mind, such as funding agencies and the public. There appears to be little consensus in the behavior genetics community about the best way to communicate behavior genetics research to the public. While some argue that scientists should engage actively in public debates about the meaning of behavior genetics results, others argue that researchers should work to keep potentially harmful formulations about the genetic basis of behaviors out of the public sphere. Neither is there any consensus on which representations of genes and behavior are "appropriately

simplified" and which are "distorted" (Hilgartner, 1990). While researchers often portray the public as too focused on genes, they also describe the public as unaware of the biological basis of many behavioral disorders. The messages that researchers at Western portray as safe to convey to the public—that behaviors arise from complex interactions between genes and environments, and that treatments will take time to develop—are similar to the arguments that they make about methodology, suggesting a connection between debates in the laboratory and debates in the public sphere.

5.1 Constructing "Behavior Genetics"

In order to communicate behavior genetics research to the public, researchers must first create a vision of the scientific endeavor that is "behavior genetics." This section examines how behavior geneticists construct representations of the past, present, and future of their field for the public. To illuminate the ways that behavior geneticists attempt to craft their public image, I draw on Gieryn's (1983) concept of "boundary work" in the sciences. Instead of treating science as a kind of activity that is easily distinguished from other activities, Gieryn examines how scientists themselves differentiate between their methods, practitioners, and institutions and other kinds of "non-scientific" endeavors. He explores several cases where the boundary between science and non-science is disputed, such as debates about the scientific status of phrenology in the nineteenth century and twentieth century disagreements about the separation of "basic science" from technical applications. In each case, Gieryn argues that scientists enhance their credibility and authority by identifying some kinds of activities as scientific and others as non-scientific. By contrasting science favorably with other non-scientific and technical practices, practitioners can "enlarge the material and symbolic resources of scientists or … defend professional autonomy" (p. 782). Gieryn argues that boundary work is especially common in instances where scientists are attempting to craft a public image for science.

Researchers in behavior genetics also engage in boundary work to differentiate their methods, facts, and practitioners from other kinds of knowledge about the heredity of behavior. I discuss three areas where boundary work is being done in shaping the public image of behavior genetics: around its connections to past histories of eugenic science, around which claims and practitioners are "scientific" and which are "ideological," and around the nature of behavior genetics as an "applied" science. Identifying the origins of contemporary behavior genetics research is a rhetorical strategy used by both practitioners and critics to make claims about the credibility of the present-day field. Behavior geneticists attempt to create a history for their field by selectively detaching from and linking to past efforts to understand the heredity of behavior, especially those closely associated with eugenic science. Behavior geneticists also manage controversial events in the recent history of their field, such as debates about the heritability of intelligence, by delineating boundaries between proper scientific practice and ideological distortions of scientific methods. Both supporters and critics of controversial behavior geneticists such as Arthur Jensen and Phillipe Rushton attempt to strengthen their positions by alleging that those who disagree with them are biased and unscientific. Finally, I explore how behavior geneticists also manage the somewhat less controversial boundary of whether their work should count as applied or basic science. Animal behavior geneticists in particular tend to portray the applications of their research to human health as quite far away, a move that is aimed at managing public expectations of the field and mitigating potential ethical concerns that might arise from using behavior genetics knowledge in clinical or policy settings.

One of the areas where boundary work is being done in behavior genetics is around what should count as part of the history of the field. Scholars in science and technology studies have pointed to some of the ways in which actors' histories can be seen as part of particular professional or political struggles. For example, in her examination of the

history of East German psychotherapy, Leuenberger (2001) compares official accounts with historical accounts told by practitioners after the fall of the Berlin wall in 1989. While official accounts exclude any mention of theories such as Freudian psychoanalysis that were seen as incompatible with socialist ideals, practitioners' retrospective accounts highlight the informal use of psychoanalytic ideas that took place prior to the collapse of state socialism. Leuenberger shows how these accounting practices can be used to stabilize psychotherapists' current professional position by emphasizing the continuity between their past and present practices.

In the field of behavior genetics, practitioners often mark the beginning of the contemporary field in 1960, the year that Fuller and Thompson's (1960) textbook *Behavior Genetics* was published, but some extend the field's history back much further. There are few attempts to tell the history of the field in the academic literature, and those accounts that do describe behavior genetics' origins tell the history of the field through the lens of the history of eugenics (Kevles, 1985; Paul, 1991) or IQ and race debates (Kamin, 1974; Tucker, 1994).[1] These accounts often connect contemporary behavior genetics research to controversial past figures such as Sir Francis Galton, a pioneer in developing statistical techniques to measure heredity (such as twin studies) and inventor of the term "eugenics." In her study of the history of eugenics, for example, Paul (1991) locates the origins of the contemporary behavior genetics field in research funded by the Rockefeller Foundation in the post-World War II period, especially selective breeding programs on the heritability of social behavior in dogs at the Jackson Laboratory. She offers evidence of the eugenic thinking present in the Rockefeller Foundation's rationales for supporting these research programs, but also some of the ways that the researchers resisted efforts to portray their findings as offering unequivocal support for the heritability of intelligence and other traits.

[1] Panofsky (2006) gives a brief history of the behavior genetics field, focusing especially on behavior genetics since 1960. He also notes the lack of historical research on behavior genetics, and argues that his contribution is a "skeleton around which a more comprehensive effort could be built" (p. 99)

Whether or not eugenics research counts as part of the history of contemporary behavior genetics is contested within the field. For those who are critical of contemporary behavior genetics, portraying twenty-first century behavior genetics research as a continuation of research programs founded by practitioners like Galton is a way of making the case for the racist and eugenic tendencies of the contemporary field. For behavior genetics practitioners, pointing to the re-emergence of behavior genetics in the 1960s is a way of placing these same events outside of the boundaries of contemporary behavior genetics. By locating the origins of their field in the post-World War II period, practitioners attempt to distance their work from the atrocities of Nazi eugenics programs, and locate their field instead in post-World War II movements to use scientific data to oppose racist ideologies in society.[2] In his study of controversy in behavior genetics, for example, Panofsky (2006) notes that many of the researchers he interviewed who were active in the early days of behavior genetics recalled that the field made explicit attempts to break with what they identified as racist and eugenicist traditions of genetic research. Panofsky points to Fuller and Thompson's seminal textbook, where the authors wrote that they debated the merits of including research from human studies on the heritability of behavior in the textbook at all since they thought so little of it met standards for rigorous scientific research (Panofsky, 2006, p. 101). A recently published introductory textbook similarly distinguishes between contemporary efforts and a contentious past by describing behavior genetics as a field with "a long past, but only a short history" (Kim, 2009, p. 3). The author disconnects the "long past" of research on the heritability of behavior—such as animal breeding efforts in ancient times and the contributions of figures like Darwin and Galton—from the "short history" of behavior genetics as it emerged in the 1960s. More recent controversies around the heritability of intelligence in different races that emerged in the early 1970s and again in the mid-1990s are subject to the same kind of boundary work. For some, researchers

[2] For a discussion of the concept of "race" in science in the post-World War II period, see Reardon, 2005.

such as Arthur Jensen and Phillipe Rushton who argue that races have different inherited capacities are anomalies or outliers in the field, but for others these researchers provide evidence of the persistence of eugenicist and racist orientations in behavior genetics.

Not all researchers advocate disconnecting from this past entirely. Practitioners' reactions to the Nuffield Council on Bioethics (2002) report on ethical issues in the study of genetics and human behavior demonstrate some of the ongoing concerns in the field around consequences of connecting to or disconnecting from past research on the genetics of behavior. The journal *Genes, Brain and Behavior* published two commentaries on the report following its publication, one of which argued that the report had gone too far in envisioning the potential ethical implications of behavior genetics research and the other arguing that it had not gone far enough (Hay, 2003; Wahlsten, 2003). Hay (2003) suggested that the report took an overly pessimistic view of the history of the field, and worried that the report would have detrimental effects on the public perception of behavior genetics work and on researchers' ability to get funding. He argued that the historical events that the working group pointed to were "so removed from what behavior genetics in 2003 is all about" and bemoaned the fact that contemporary behavior genetics can't seem to "move beyond a tainted history that was not of its own making" (p. 321–322). While Hay's opinion piece argued that the report had not done enough to distinguish behavior genetics research today from the histories of eugenics and racist science, Douglas Wahlsten, an animal behavior geneticist from Canada, argued that the report elided some of the links between eugenics and contemporary research. He charged the panel with "airbrushing" the concept of heritability in the report, writing:

> The historical review of eugenics in relation to behavior genetics and heritability analysis does not fully confront past or present realities in our field. The Nuffield document acknowledges gross abuses of genetic theory in the 1920s and 1930s but denies that this sordid past means that "contemporary research on the genetics of behavior is in any sense eugenic or is driven by considerations that could be considered eugenic" (Nuffield Council on Bioethics 2002, p. 22). While noting that there has been controversy in recent

> times concerning the heritability of intelligence in particular, it fails to make a clear connection between the concept of heritability and eugenical selective breeding. On the contrary, it suggests there is no connection in contemporary science. (Wahlsten, 2003, p. 327)

Wahlsten continued on to fill in some of the needed history as he saw it; namely, that the concept of heritability is intimately tied to histories of selective breeding. He pointed out that heritability measures were originally developed for the purpose of selectively breeding farm animals, and that contemporary behavior genetics researchers such as Aurthur Jensen and Richard Herrnstein, who argue for the high heritability of intelligence, also make direct links between heritability and social policies that might be described as eugenic. Wahlsten's article connects the concept of heritability estimates to different histories than those mentioned in the report. By denying the sharp boundary that Hay draws between contemporary work on heritability and these past histories, Wahlsten also seeks to warn practitioners and the public about the potential for present-day eugenic applications of behavior genetics concepts.

Researchers at Western also selectively connect and disconnect past histories to contemporary behavior genetics in the way that they present the field for new practitioners. In the introductory behavior genetics class, Dennis used one of Wahlsten's (1994) articles critiquing the concept of "heritability" on the syllabus. The article argued that human studies that simply "partition the variance" of a particular trait into genetic and environmental contribution are fundamentally flawed because they do not consider the interaction of genetic and environmental factors, and that these heritability estimates are of limited value to the field or to human health anyway. Wahlsten argued that researchers should instead focus on trying to understand the mechanisms of gene action through animal models and gene-environment interaction effects. The students were surprised by the tone and force of this critique in an otherwise rather dry set of articles on heritability. Dennis explained that part of the background to this article was that Wahlsten was very involved in researching the history of eugenics in Alberta and was concerned with the direct effects

of how heritability estimates could be misused (see, for example, Wahlsten, 1994).

Another boundary that is important in constituting the field of behavior genetics is the boundary between what counts as "science" and what is considered "ideology." Some might argue, for example, that Wahlsten's attempts to expose flaws in heritability studies because of his concerns about the possible eugenic applications of these concepts are the actions of a socially responsible scientist, while for others it might be evidence of how left-wing ideologies can color a scientist's judgement. As Reardon (2005) has explored in her study of the Human Genome Diversity Project (HGDP), commentaries on scientific investigations of racial differences have often portrayed science and ideology as separate and opposed entities, where objective and neutral scientific approaches produce truth and ideological influences distort this process. Reardon describes post-World War II efforts to redefine "race" as a social category without a biological basis as attempts to shore up the credibility of the biological sciences by ensuring that science would not be distorted by social biases as it had been so horrifically during World War II. Reardon argues, however, that science and society can never truly be decoupled: She carefully chronicles how the concept of "race" never really disappeared from science, and how HGDP researchers' attempts to treat the scientific investigation of race as separate from its social meanings in other contexts incited forceful opposition from other researchers and interest groups.

Despite the deep problems that Reardon and other science and technology studies scholars have identified in treating science and society as separate categories, the formulation of science as opposed to ideology provides powerful rhetorical resources for actors to discredit particular scientists or scientific approaches.[3] In their discussion of the scientific community's response to accusations of fraud in Sir Cyril Burt's research, Gieryn and Figert (1986) demonstrate how these demarcations were used to discredit Burt and protect the authority and credibility of the field. Cyril Burt was a British psychologist well-known

[3] See also Gerovitch, 2002 for a discussion of the ways in which practitioners selectively aligned and distinguished cybernetic science from ideology in the former Soviet Union.

for his studies on the inheritance of intelligence; but following his death in 1971 he was accused of falsifying data and inventing fictional co-authors for his papers. Gieryn and Figert describe psychologists' reactions to the unfolding controversy around Burt and his data as a kind of "status degradation ceremony" (Garfinkel, 1956) that transformed his status from that of a respected scientist to a mere fabricator of data. Descriptions of Burt's tendencies to be hasty in drawing conclusions, to ignore conflicting evidence, and most importantly, his willingness to fabricate data to support his own position on the heritability of intelligence placed him outside of the boundaries of proper scientific practice.[4]

The behavior genetics community's responses to Glayde Whitney's presidential speech at the 1995 meeting of the Behavior Genetics Association demonstrates how researchers also use the boundary between science and ideology to discredit particular practitioners and distance themselves from particular visions of how behavior genetics methods should be used. At the closing banquet celebrating the 25th anniversary of the society, Whitney argued that it was time for the field to move towards a new agenda of studying differences between human racial groups. He presented data on the murder rates in countries and cities with different "racial compositions" in their populations, and argued that it was a reasonable hypothesis to investigate whether these differences in the murder rates could be attributed to genetic differences between races. According to the recollections of those present during the banquet, many of the members of the audience were shocked and surprised by the speech, and embarrassed that the mostly black waitstaff had to listen to it (Panofsky, 2006, p. 147–148). Some members of the audience, including a few seated at the executive table, walked out of the banquet in protest.

Whitney's speech and the subsequent reactions to it show how the behavior genetics community used science and ideology to create demarcations in the field and locate

[4] See also Gould, 1981 for an extended critique of Burt's work.

some practitioners on the wrong side of these boundaries. In a preface to a transcript of the speech that was published in *Mankind Quarterly*, Whitney wrote that his speech was motivated by a desire to break the "ideologically-based dogma and taboo" against investigation into the genetic basis of racial group differences (1995, p. 327). In his speech, he pre-emptively accused those who might disagree with his vision for the future of behavior genetics of having "Marxitis." He explained:

> Many of the scholars that suffer from Marxitis do not realize that they are infected. The symptoms of this disease include an intellectual bias, an insistence on environmental determinism as the acceptable cause of group differences. In severe cases, it includes an unbending intellectual absolutism akin to medieval scholasticism. It is lethal to honest science.

Whitney contrasts the "honest science" that he is advocating for to the "unbending intellectual absolutism" of the "Marxist" critics who might disagree with his new agenda. In Whitney's portrayal, his approach is objective and scientific, and his detractors' commitment to the idea that there are environmental explanations behind racial differences is nothing more than dogmatic liberalism.

Behavior geneticists who were critical of Whitney's position also inscribed a boundary between science and ideology in their analysis of these events, but located Whitney on the side of ideology. Many of the behavior geneticists that I interacted with, including Dennis, described the speech as blatantly "racist." Jack, a postdoc at a university on the West Coast and a former student of Dennis's, told me that he uses a transcript of the speech in one of his courses to illustrate the potential for racist misuses of behavior genetics research. Other graduate students at Western, who were not present when the speech was delivered, described the incident to me as a "bell curve like speech," referencing Herrnstein and Murray's (1994) controversial bestselling book *The Bell Curve,* whose claims about the heritability of intelligence and differences between races generated heated debate in behavior genetics community and the popular press. One *New York Times* columnist, for example, described the book as an attempt to "drape the cloak of respectability over

the obscene and long-discredited views of the world's most rabid racists" (Herbert, 1994). These criticisms of Whitney (and Herrnstein and Murray) argue that their positions that are ideologically motivated, and that they are using the guise of science to promote racist agendas.

Finally, another boundary that researchers negotiate in the public presentation of their work is how quickly behavior genetics research will lead to new applications that can be used in clinical settings. Animal behavior geneticists in particular describe their field as occupying a peculiar position between basic and applied science. Researchers at Western, for example, receive much of their funding from institutes at the National Institutes of Health and told me that they would probably not be able to secure funding from the National Science Foundation because of the health-relevant focus of their research, but at the same time many identify their research as "basic science" that is far from producing any applications. By identifying animal behavior genetics as "basic" research and emphasizing the long time lag between current research and clinical applications, practitioners manage the perceived distance between animal research and human applications for funders and the public (as well as for each other).

Many researchers at Western talk about the future applications of their research in quite general terms, and usually in the context of grant writing, public communication, or describing their motivations for entering the field. In some cases, researchers resisted the idea that they should have to connect their research directly to human health at all, especially when the ways that their research might "translate" to the clinic seemed difficult to identify. When I interviewed Alex, a graduate student at Western, I was surprised to find that he said very little about the clinical relevance of his work, especially since the description of his laboratory's research on their website referenced the importance of "translational" research several times. Even when I asked him directly about the applicability of his research to human health, he seemed quite reluctant to talk about it:

> *Alex:* How close to addiction is that? I mean, it's not close at all. Yeah, certainly there's some brain changes going on, but do they match the changes that would have gone on if I'd given them a thousand injections, or that the humans who do cocaine every day for 17 years? No, it doesn't. And I don't, honestly I don't tell people about my research, I vaguely, I rarely say the word addiction unless it's someone who doesn't know about the research, because that's an easy term to relate … You couldn't really say that it's related to addiction at all, other than to say "hey, maybe if we can do that in rats, maybe the same pharmacotherapies might have the same effect in humans."
>
> *NN:* Yeah, well, I was going to mention that.
>
> *Alex:* At some point, in about 74 years.
>
> [laughter]

Alex strongly resists the idea that his work is directly related to human addiction or that it has direct relevance for human treatment. Talking about his work as related to "addiction" is a simplification used only when talking to outsiders about his research, and he sees himself as a basic scientist studying learning and memory processes in animal models. He suggests that there might be some link between the drugs that change addiction behaviors in rats and drugs that could also work in humans, but he portrays the potential application of these results as something that would take place in the distant future—74 years into the future, to be precise.

I argued in chapter 1 that this depiction of animal behavior genetics as a "young" field that is far from producing definitive statements or future applications is strongly tied to an understanding of behavioral disorders as "complex disorders." But even in instances where researchers drew on different models of gene action, they often still located potential clinical applications of their research in the distant future. Thomas, a researcher in Germany working with animal models of anxiety, also works in a setting with a strong "translational" focus. His laboratory is affiliated with a clinical center for anxiety treatment that is located in the same building to allow for interaction between researchers investigating anxiety and physicians who are treating it. Thomas's description of gene action was notable because he was the only researcher in my study to use a "disease gene" construction in an interview.

He says:

> If you marry by chance a man who has the same constellation, that means an anxiety gene that is suppressed by a dominant non-anxiety gene, then your kids could have alleles, and then these alleles, the two repressed genes become dominant. And this child then will suffer from an anxiety disorder sooner or later regardless of what the environment is, what the environment looks like in which the boy or the girl grows up. This is the genetic predisposition, and it will work. The disease will appear sooner or later, maybe facilitated earlier by a stressful environment, later in a very peaceful and stress-free environment, but it will become evident sooner or later.

Far from presenting a picture of the etiology of behavioral disorders where gene action is nearly hopelessly entangled with the effects of other genes and environmental factors, Thomas gives a description of the genetics of anxiety that is classically Mendelian and strongly deterministic. He describes a fictional scenario where a child might inherit inheriting two recessive "anxiety genes" from her parents that eventually lead to an anxiety disorder "regardless of what the environment is." But even though his manner of describing gene action is quite different from how researchers at Western typically talk about genetics, he also portrays clinical applications as far away. In the course of the interview, he went on to complicate his description of "anxiety genes" by suggesting that there are likely to be many anxiety genes, each of which has a small effect:

> *Thomas:* And that might explain why we don't have the wow, the breakthrough in a sense that somebody discovered the magic bullet. There is none. And that's the problem. So it is step by step, this is a five percent contributor, okay, it's accepted by the scientific community, so let's go to the next. Then if we found twenty five percent contributors, then we are close to a hundred percent, and then we can try to characterize patients in a sense that we can design a cocktail that is most promising for this particular person.
>
> *NN:* So it will probably be more of an incremental breakthrough where you'll slowly find more genes that influence it and maybe drugs that act on those systems.
>
> *Thomas:* Yeah, at least in mice. And then we still have to look for homologies in the clinic. This is another long way to go, but there's agreement that we should start with mice.

Thomas doesn't specify who holds the assumption that there should be a magic bullet to cure anxiety—the public? funding agencies? pharmaceutical companies?—but from his description it is presumably not scientists. He argues that the scientific community is in agreement that there are no magic bullets, and that the most logical course of action is to proceed by identifying and verifying many genes that each have a small influence on anxiety phenotypes using animal models. When I attempt to reflect this vision back to him for confirmation, suggesting that the field will advance through a series of incremental discoveries that will lead new pharmacotherapies, he inserts even more distance into my timeline by pointing out that there is still a level of translation that needs to happen between the mouse and the human even after promising gene candidates are identified. Although the distance between current research efforts and these imagined futures varies in scientists' descriptions, the animal behavior geneticists that I interacted with almost universally portrayed clinical applications as difficult to accomplish and quite far away, pushing back on perceived expectations from funding agencies or perhaps a lurking public that their research will produce drugs that can cure anxiety.

To summarize, the public image of the behavior genetics field is one that practitioners work to define against several contested boundaries. Whether contemporary behavior genetics research on topics such as intelligence and race is a continuation of eugenic science or not, or whether it is motivated by scientific curiosity or ideology, are questions that are debated both within the field of behavior genetics and in much more public venues. These boundary disputes have developed into highly publicized and highly controversial disputes at particular moments in the field's history. In other cases, researchers define their work against less identifiable and less adversarial assumptions about the field, such as the idea that the field's work will produce "magic bullets" to cure human behavioral disorders. Through their engagement with these boundary disputes, researchers construct an image of "behavior genetics" as a field that is disconnected from controversial past

histories of eugenic research (although some warn that practitioners need to be on guard against those who might reconnect it to eugenic social policies), that is defined by scientific approaches to studying behavior and not ideologically motivated inquiries into human group differences, and that has an applied orientation but is still distant from any future applications in clinical settings or science policy.

5.2 Constructing "The Public"

The public communication of behavior genetics research not only requires an understanding of what "behavior genetics" is as a scientific endeavor, it also requires an understanding of who "the public" is that this scientific information will be communicated to. Science communication scholars have argued that "the public" is not a monolithic entity that responds with one voice to scientific and technological developments, nor does the public speak for itself. Surveys, focus groups, citizen's panels and consensus conferences designed to measure public opinion on science and technology each constitute the public in different ways; gathering opinions from "representative" publics, "interested" publics, or "knowledgeable" publics. Wynne (1991) has argued that one of the dominant ways of conceptualizing public communication of science is what he terms the "deficit" model, where the public is seen as lacking information about science and providing education to the public is seen as the solution to garnering increased support for new research and technologies. Wynne argues that this way of conceptualizing the public obscures the knowledge and experience that laypeople might already have about scientific controversies, as well as the range of reasons they have have for objecting to particular technologies. Recent research has pointed to the role that scientists themselves play in generating "imagined lay publics" (Maranta, Guggenheim, Gisler, & Pohl, 2003) with different types of attributes, such as publics that need to be informed of scientific facts, publics that can be engaged in discussions, or publics that are attentive followers if not always enthusiastic

supporters (see, for example, S. R. Davies, 2008; Michael & Brown, 2000; Stilgoe, 2007; Barnett, Burningham, Walker, & Cass, 2010).

This section examines how behavior geneticists construct representations of "the public" or "publics" that respond to behavior genetics research and of the political culture that researchers work in. In contrast to their laboratory work, where researchers have highly specialized methods and techniques to produce knowledge about the genetics of behavior, here scientists make sense of the public in "ordinary" ways (Garfinkel, 1967). As Michael and Brown (2000) put it, scientists are engaged in a process of "lay political science" where they make their own observations and assumptions about the nature of political processes, especially as they pertain to the interaction between scientific experts and the lay public. Scientists draw on diverse sources of information around them, such as interactions with friends and family members, experiences with reporters, popular books and television shows, and public debates about science. They do not necessarily draw on these sources in any systematic way, and the types of information that researchers at Western draw on to make sense of the public are often highly specific to their cultural context and their position in the behavior genetics field. When talking about how the public perceives behavior genetics, researchers at Western often discuss debates about animal rights activism and the heritability of intelligence that reflect their position as animal researchers and behavior geneticists with psychology backgrounds. In other cases, researchers draw on public controversies about science to make sense of their political culture that are unique to the American context, such as debates about intelligent design and global warming. I explore how researchers use these contentious and divisive debates to understand the political culture that they work in and the publics that they are responding to.

One of the ways that behavior geneticists make sense of their political culture is by drawing on historical instances where behavior genetics methods and results have been the focus of public discussion. Many of the senior animal behavior geneticists that I interacted

with began their careers in behavior genetics in the late 1960s and early 1970s, at a time when questions about the relationship between genes and intelligence in particular were highly politicized in the United States. Senior researchers often pointed to the publication of Arthur Jensen's (1969) article "How much can we boost IQ and achievement?" and the ensuing public controversy over the article's claims as a formative moment in their careers. In a lengthy article reviewing data from both human and animal studies on intelligence, Jensen argued that IQ scores are a good measure of intelligence, and that intelligence is a characteristic that is highly heritable. He went on to conclude that educational programs designed to increase IQ scores in below average students were destined to fail since environmental factors only accounted for a small portion of the variation in intelligence scores. In particular, he argued that efforts like the "Head Start" program funded by the United States government to increase scholastic achievement in black youth were unlikely to succeed because intelligence was largely determined by genetics.

The backlash against the publication was immediate and forceful, provoking widespread criticism from inside and outside the scientific community and protests on university campuses. In his study of controversy in the field of behavior genetics, Panofsky (2006) writes:

> It is perhaps difficult to imagine the sense of threat and danger that Behavior Geneticists felt during this period. This was not merely a heated intellectual disagreement with nasty invective exchanged in the pages of academic journals. Protesters disrupted their academic meetings, fist-fights broke out, and scientists received threats to their families. … Behavior Geneticists felt that the very existence of their field was in jeopardy as a result of the IQ controversy set off by Jensen's work. Whatever their position on the ideas, Behavior Geneticists had the palpable sense that they were fighting for the field's life (p. 112–113).

The dramatic nature of these events left lasting impressions on many researchers in the field, especially for those who were practicing behavior geneticists at the time when the article was published. Anthony, a senior researcher at a university in Canada, recalled

that the publication of Jensen's article transformed his understanding of his work and his identity as a researcher. He says:

> I did a postdoc at the Institute for Behavioral Genetics at Boulder, Colorado, where they work with mice. So I began to learn about inbred strains and behavioral testing in mice because I had just done dogs before, so it was a lot to learn, a new species. Well, I'd only been there a couple months when—and like I say, I was really more of a pure scientist, interested in the biological basis of learning and memory—and all of a sudden, boom! This bomb was dropped into the middle of our work called the *Harvard Educational Review*, an article written by Arthur Jensen on environment, genetics, and intelligence.

The experience of responding to criticisms of the field during the race and intelligence controversies of the 1970s still impacts the way that some behavior geneticists approach the public communication of their research. On the first occasion that I interviewed Anthony, he cut short my introductory questions about his career history and background as a researcher, and told me that he had participated in studies like mine before. He asked if I "wanted to know how he got into this race and intelligence stuff," and then stood up to retrieve a copy of the issue of the *Harvard Educational Review* containing Jensen's article from a bookshelf near his desk while explaining to me why he felt a responsibility to be knowledgeable about intelligence, race, and heredity following the publication of this work. Scott, another senior behavior geneticist who was working in the field at the time Jensen's article was published, explained to me that the controversies of the 1970s also shaped the way he thought about the public promotion of his work. He says:

> Let me give you another little bit of insight about my own particular reticence with regard to the media and promoting science. A lot of it has to do with the fact that it's behavioral genetics. You may or may not be aware of the large-scale controversy about genetics and intelligence that flared in the late 60s and early 70s. There were several animal behavior geneticists who were the targets of protests, and it was ugly at [my university] for a while. Anyone having anything to do with behavioral genetics was automatically labeled a racist, and so there's a sensitivity about the way that that information is imparted to the general public that a lot of us still have.

Scott's and Anthony's responses to my request for interviews show how researchers—especially senior behavior genetics researchers—make sense of the political culture of

the present through their experiences of the political culture of the 1960s and 1970s. They imagine that behavior genetics facts could still become controversial in a similar way, and in some cases pre-emptively respond to critiques of their field's research that they think might be raised. Herrnstein and Murray's (1994) book the *The Bell Curve*, which also argued for a racial difference in intelligence, and J. Philippe Rushton's (1995) book *Race, Evolution, and Behavior*, that compared and ranked physical and behavioral characteristics of different racial groups, renewed these debates in the mid-1990s, providing more evidence for behavior geneticists' understandings of their subject matter as one that could (and probably would) become contentious again. As one researcher put it to me, public controversies about race, intelligence and, genetics just seem to "come around every decade or two" in the behavior genetics field.

Understanding the political culture of the present through past debates about race and intelligence makes the terrain on which behavior genetics facts might travel seem quite hazardous. Scott elaborates on why it is that he is hesitant to talk to the media about his work or anything having to do with behavior genetics:

> All those things that have social implications that can immediately start … You can immediately start relating them to racial and group differences in humans, and that's dangerous ground. Even if you come to the material in a completely objective way, there are going to be attempts to distort what you have to say, and draw conclusions beyond what the data would permit.

The "imagined laypersons" that are responding to behavior genetics research in Scott's view are not only unaware of the limitations of behavior genetics research, but also have the potential to misuse behavior genetics facts. Who exactly is attempting to "distort" behavior genetics data or to "draw conclusions beyond what the data might permit" is not entirely clear from his comments, but other practitioners identify many different social groups—such as journalists, members of the public, "right wing reactionaries," or even members of their own field—as potential culprits.

Another public debate that is highly salient for researchers at Western is the debate

about the use of animals in research. Mouse researchers are not typically the direct targets of protests or campaigns, but researchers often discuss the way in which the climate of animal rights activism nonetheless impacts their personal lives and the way that they talk about their research outside of the laboratory. Researchers say that discussing their research to "random strangers" is especially difficult in Western City because animal rights groups like People for the Ethical Treatment of Animals (PETA) are particularly active on the West Coast of the United States, and they never know if they might be talking to someone who is especially sensitive about the use of animals in research. One grad student at Western, for example, recounted that when she adopted a cat from a local animal shelter she made the "mistake" of mentioning that she was a researcher at Western University to the shelter employee, who made her sign a form saying that she would not use the cat for animal research. The grad student was insulted and dismayed that the shelter employee could possibly believe that she would go through a lengthy adoption process only to use the cat for research. On another occasion I was out with Alexis, a technician in the Smith laboratory, when a friend asked her about how things were going at work. Alexis told her that she was learning to do brain cannulations in mice, a surgical procedure where a small tube is inserted into a precise location into the mouse's brain so that researchers can later deliver drugs to a particular brain region or sample brain fluid to measure neurotransmitter levels. This type of surgery is tricky to perform and Alexis was excited that Sharon had asked her to be the one to learn this technique. Her friend was clearly not as impressed, and the conversation ended quickly and uncomfortably. The impacts of animal rights activism on the personal lives of scientists can be profound for researchers who are working with other kinds of animal models, especially non-human primates. Susan, a researcher at a university on the West Coast who works with monkey models, was deeply personally impacted by anti-animal research movements. After receiving threats from animal rights activists, she now changes her license plates frequently, wore a

"disguise" with a wig for her picture on her university's website, and her daughter was even placed under FBI protection for a period of time.

During my time at Western University, there were several small-scale protests against the use of animals in research that took place on the university campus. These protests, which seemed to be somewhat regular occurrences, often provided a starting point for conversations about the public communication of behavior genetics research in the laboratory.[5] Chloe, reflecting on one of the recent protests in an interview with me, expressed sympathy for the activists' position:

> I think as researchers you do have to be sensitive to people who might be a little bit more averse to using animals in their research. And you know, that's been a challenge. That's something that's very important to me is to minimize the pain and the trauma to the animal for research. And it's something that's just … I, I hate, hate sacrificing animals. I mean really. It hurts, it hurts me. And I don't know everybody thinks like that, especially a lot of the activists, you know the protestors … I mean nobody asked [the animals] to, you know, give their lives for research. We breed them and we raise them and things like that. And so you try to justify it to yourself, using an animal, but it's still really hard. But that's kind of something, you know, you try to be a little more sensitive. Like with my sister now, because I didn't realize that she felt that way.

While Chloe says that she can understand why animal rights activists would be opposed to the pain caused to animals in the course of research, many other researchers expressed anger about what they felt to be misrepresentations of their research on the part of activist groups. On the day of one protest, the lunchroom conversation in the Martin laboratory was abuzz with discussions about animal rights activism. The majority of the laboratory members were of the opinion that animal rights protestors didn't really know what research was actually like, and that (as Chloe's comment reflects) researchers don't enjoy killing animals either. Ava, who was browsing the internet while eating her lunch, commented

[5] Protests that would be taking place on the campus were announced via e-mail and in lab meetings, and the members of the Smith laboratory seemed well-versed in how to deal with these events. Several people, for example, explained to me what areas of the campus protesters could legally picket in, and routes to take to the laboratory that avoided entrances that were likely to be picketed.

that she was researching some arguments she could use to rebut the claims of activists who refused to acknowledge any useful role for animals in research, and said that she was considering going downstairs to argue with them later on. Laura, who walked into the lunchroom in the middle of this conversation, commented that she tried to use those moments as "teaching opportunities" to talk about the benefits of the kind of research that they do as a laboratory.

The voice of animal rights activists is another one of the "publics" that researchers imagine they are responding to when they talk about animal behavior genetics in public settings. Even when criticisms about the use of animals in research are not directed at their work in particular, practitioners say that they feel as though they should speak out on behalf of animal research as a whole and address laypeople who may be swayed by activists' arguments that animal models do not contribute to understanding human health. Dennis, for example, said that he was enthusiastic about putting out a press release on some of the laboratory's work with the ARC models for binge drinking because he thought that particular project would provide a convincing public demonstration of the benefits of animal research. He explained:

> You can say, without dissing the existing models, you can emphasize that this a better model because mice really do drink and fall over just like college students. So that's cute enough and informative enough that we'll probably try to get some mileage out of that. I mean at this point, I actually don't need publicity. … But that message I think is really important to get out there, because we're in kind of a long scale war with PETA and the rest of the animal rights movement over the rationale and the validity of animal models at all. And you know, I could easily hide out here because nobody cares about rats and mice. I'm not using kittens or monkeys or dogs, something cute that people really get upset about. But it's really important to counteract that, and scientists more and more realize that we're doing the whole field a disservice by keeping our heads down on that score.

Dennis portrays the public dissemination of his work with animal models as a way of engaging the claims of groups like PETA that animal models are not valuable. Researchers at Western also engage animal rights activists as imagined adversaries in more subtle ways

in their day-to-day work in the laboratory. For example, a poster depicting two white rats with the caption "They've saved more lives than 911" hangs in the hallway next to Dennis's office, providing a pro-animal research message to anyone who steps off of the elevator.

Liam, a graduate student in the department, says that he often pre-emptively defends himself and his research against animal rights critiques when he talks about his work publicly, even when no one is directly asking him for such a justification. He says:

> I find myself, even to my family, trying to justify the research I do, and I feel like it's necessary even if, I don't know, even if they aren't really asking for that. Maybe I have a guilty conscience or something about it. But I mean, I shock mice. That's a very nice thing to do to another living creature. And so trying to justify why I do that is something that I do struggle with some on my own.

The presence of animal rights activism in the imagined terrain of the public creates a kind of "approach-avoidance conflict" for researchers: They want to provide examples about the positive contributions of behavior genetics research and work with animal models in particular, but at the same time they are fearful of overstating the benefits of behavior genetics research or risking a run-in with an animal rights person in social settings.

Researchers also use opinions on the use of animals in research as an indicator for how particular individuals that they interact with might respond to animal behavior genetics research. In their examination of how animal researchers conceptualize the public in animal rights debates, Michael and Birke (1994) point to the ways that researchers disaggregate the public into groups with different characteristics, such as those that debate the issue "rationally" rather than emotionally, or those who are more and less amenable to researchers' arguments. They argue that these demarcations are a way of managing public debates by restricting those who can legitimately participate in them. At Western, researchers' portrayals of the activist segment of the population as dogmatic and irrational contributes to the perception that it is impossible to be against the use of animals but supportive of behavior genetics. Dennis, for example, suggested to me that it was better to think of the animal rights movement as a "religion," a framing that portrays them both

irrational and steadfast in their convictions.

Researchers also use attitudes towards animals to parse the lay public into segments that are more or less likely to be favorably disposed to their research or science in general. When I asked Ava whether it was easy to talk to her friends and family about her work in the laboratory, she uses sympathies with the animal rights debates as a shorthand for explaining why some of her relatives are interested in her research and others aren't. She says:

> Um, it depends on what their background is. Like you know my sister, she's like a big animal lover. She probably should have been a vet but she never did, I don't know why. But I know she's kind of grossed out by it, but she doesn't say anything to me about it. But then, like my boyfriend's parents, one of his brothers is a recovering meth addict, and so they are, they love it. They love hearing about it and they love knowing what's going on, because it's directly impacted their family, you know? And so they believe in it. And so, you know, I think that makes things different.

Her sister's status as an "animal lover" becomes a proxy for assessing how her sister might value the animal behavior genetics endeavor as a whole, while her boyfriend's family's personal stake in addiction issues is offered as an explanation for their positive reactions to behavior genetics research.

Finally, researchers also draw on public debates about the status of scientific knowledge in general (rather than behavior genetics in particular) to construct an image of the public that they are communicating with. Debates about the evidence for global warming, the status of evolution and intelligent design as scientific theories, and the Bush presidency were a few of the elements of American culture that researchers at Western used to talk about the anti-science aspects of their political culture. For example, in a discussion with Dennis about his impressions of the public's reaction to his work with animal models of addiction, he used public debates about the status of evolution as a proxy for assessing how the public might respond to his own research:

> *NN:* So do you think then that the general public is unconvinced of the idea that you can use animals to model [drinking behaviors]?

> *Dennis:* Let's not forget that more than 50% of the general public believes that you and I aren't biologically related to great apes. [laughter] Right?
>
> *NN:* Yeah, okay, fair enough.
>
> *Dennis:* Fifty percent of the public does not accept the basic findings of general evolution. So, yes. I mean there are lots and lots of folks out there, I don't know what percentage, that are just inclined to discount animal research as being not a good thing to do or not valuable to do. But most of the general public doesn't understand research at all. Doesn't understand science at all, because as a profession we've done such a terrible job at explaining what science is, and what research is and how it relates to the rest of life, you know? That's a general educational failure.

To Dennis, debates about the scientific status of evolutionary theory suggest a general lack of scientific knowledge in the American public, and consequently he suggests that it is unlikely that this public will understand or support animal behavior genetics research. Researchers at Western also reference debates about science in American culture to talk about the role of the scientist in managing public controversies about science. As Dennis's comments above suggest, some researchers see the public's lack of understanding of science (as evidenced through debates on evolution and global warming) as a failure on the part of the scientific community in communicating to the public. For new scientists at Western, the first day of the introductory behavior genetics class began with a lecture about evolutionary theory given by Ruth, who paused at several points to emphasize that it was the responsibility of scientists to demonstrate that there was substantial agreement about the status of evolution as a scientific theory and not to over-emphasize the extent of controversy within the scientific community. As I walked back towards the laboratory buildings with her after class, she explained that she thought evolutionary theory was foundational to the research that she did in behavioral neuroscience, and was disturbed that some members of the scientific community contributed to ongoing debates about its status as a scientific theory. She pointed out that one of the faculty members at Western University was an outspoken creationist. During a discussion about the relationship between science and politics in an interview with Ava, she commented that one of the past

candidates for the graduate program held similarly controversial views about the science behind global warming that did not receive a positive response in the department. She said:

> *Ava:* Or like … you know, the rare student, like one year we had a student applicant who interviewed here, who described *An Inconvenient Truth*, you know the Al Gore movie, as Al Gore's propaganda machine. And so he didn't get accepted. [laughter] You know, stuff like that.
>
> *NN:* So he didn't believe in global warming, is that what he was implying?
>
> *Ava:* Um, yeah. He did in later conversations. I think he probably did actually mean he thought it was Al Gore's propaganda machine, but yeah, he didn't believe in global warming. So yeah, that stuff.

It is unlikely that a prospective graduate student was literally not admitted to the program because of his views on global warming, but Ava's obvious distaste for the student's provocative political opinions suggests that his stance on global warming was not compatible with how Western researchers envision the role of the scientist in public controversies.

To summarize, the resources that researchers at Western use to make sense of the political culture around them and to imagine the publics that might respond to behavior genetics facts are typically polarized and highly contentious debates, such as the debates about the heritability of intelligence and its relationship to race in the 1970s, debates about the use of animals in science, and debates about the scientific status of evolution and global warming. The imagined publics that are crafted from these available resources are full of imagined adversaries, and the terrain on which behavior genetics facts might travel seems more treacherous than the terrain envisioned in the "deficit model" of science communication. The "imagined lay persons" reflected in animal behavior geneticists' discussions of public communication are not only deficient in their knowledge of the scientific method and behavior genetics facts, but are unlikely to be swayed by the provision of additional information. Scientists' descriptions of animal rights activism as a "religion" suggests that they envision themselves as engaged in protracted and unwinnable debates with certain

segments of the public that are capable of actively misinterpreting and misrepresenting the facts that are given to them.

5.3 Communicating Behavior Genetics on Treacherous Terrain

The contentious and divisive debates that structure both researchers' presentations of the behavior genetics field and understandings of the publics that they communicate to create a model where the public communication of behavior genetics facts seems incredibly difficult. By their own descriptions, behavior geneticists are surrounded by dilemmas on all sides: They want to promote their research in order to secure public support and funding and to counter animal rights arguments, and they worry that if they do not promote their work at all then they risk "sinking without bubbles," as one researcher put it. But at the same time, they fear that their work will be seen as an extension of the eugenics movement, that the public might distort or misrepresent behavior genetics findings, or that their research will be used to support ideological positions that they believe are dangerous or harmful. How do behavior geneticists map out courses of action and decide what kinds of facts can travel safely on this treacherous terrain?

This section explores how researchers conceptualize their roles in managing the public life of behavior genetics facts. I argue that there is no consensus in the research community on what stance practitioners should take in public discussions about behavior genetics, or on what kinds of messages about genetic contributions to behavior can be safely conveyed to the public. Researchers adopt a wide variety of personal positions on the public communication of behavior genetics research, and they also identify many different popular accounts about genes and behavior and about the future promise of behavior genetics research as "distorted" rather than "appropriately simplified" (Hilgartner, 1990). I argue that researchers attempt to manage these "distorted" representations of behavior genetics

research using narratives that are similar to those that they use in methodological debates in the scientific community—they emphasize the "complexity" of behavioral disorders and the long distance to applications of behavior genetics research.

Faced with the challenge of communicating what they see as a contentious subject matter to a public that misunderstands and misrepresents much of the information that it receives, behavior genetics researchers adopt many different positions on how to best disseminate behavior genetics facts. Panofsky (2006) argues that the race and intelligence controversies of the 1970s united many behavior geneticists in defense of both the emerging research field and Jensen. While not all agreed with Jensen's arguments on the heritability of intelligence, many thought that the ideas should be taken seriously, and discussed and debated within the field. In his interviews with behavior geneticists, Panofsky found that many members of the field, especially human behavior geneticists, tended to react with embarrassment and some degree of sympathy towards contemporary researchers who made provocative claims such as Rushton, but vilified critics from "outside" of behavior genetics who hampered the field's progress while contributing nothing to it.

In contrast, Anthony recalls that his experience with these same controversies led him to conclude that it was his responsibility as a scientist to engage actively in these public discussions, and not in support of Jensen. He says:

> Frankly, I wasn't that keen on having my chosen field of study be an instrument for reactionary right-wing political agendas. So I felt a social responsibility to look into [genetics and intelligence research]. And I've done that since then, I've made a point of reading books about heredity and intelligence, pro and con. And I've written a number of chapters and articles on that topic myself, and sort of joined the fray. I feel a social responsibility to do that as a scientist, that's what it amounts to.

To maintain the scientific credibility of the field and prevent it from being used to support racist social policies, Anthony himself authored a number of articles critiquing studies on the heritability of human behaviors both for scientific and popular audiences.

Scott's response to these same events was to attempt to carefully control the way that

his research was released into the public sphere, often by engaging with the media as little as possible. His concerns about the potential misuses of behavior genetics facts have led him to be extremely cautious about the public promotion of his work, and throughout his career he has avoided producing "attention seeking" press releases and media attention that he believes some other behavior genetics labs generate. During one interview I mentioned media coverage of a recently published study on depression that described the knockout mouse that they tested as a "permanently cheerful" mouse ("Ever-happy Mice May Hold Key To New Treatment Of Depression", 2006). Scott commented that this was exactly the type of statement that he wanted to keep out of the popular press:

> *Scott:* I reel from that stuff, Nicole. I think it's very harmful. And students in my behavior genetics class, early in the semester, they're infuriated by this. They want simple stories.
>
> *NN:* Infuriated by your approach?
>
> *Scott:* Yeah. They want really simple stories. But that's what they're accustomed to getting from their array of other undergraduate classes. This does this, this does this. … I can't bring myself to oversimplify things that way. So, I teach the class from that point of view but ultimately there's a complexity that we can only really scratch the surface on for most of the models that we use with animals.

Even in his undergraduate teaching, Scott feels that he has to be very careful about the kinds of statements that he makes about the heritability of behavior and the utility of animal models. For Scott, this means refusing to produce the "simple" stories and statements about behavior that he thinks undergraduates and the media want to hear.

Finally, some researchers offered more optimistic positions on how behavior geneticists could help direct the way that their research is interpreted by the public and incorporated into policy. Matthew, a graduate student who often wore a t-shirt from the NIH with the slogan "Ask me how I spend your tax dollars" around the lab, argued that behavior geneticists could help educate the public and policy makers about the potential uses and misuses of their research. He identifies several concerns about potential applications

of behavior genetics research, and argues that education is key to avoiding these future scenarios:

> You talk about eugenics, that's a big one. You talk about sort of selective eugenics, like for childbirth. This embryo has something that we don't like, so I'm going to toss it. That can certainly be an issue. I don't know, there's a million of them, so … definitely concerns. But at the same time I think as long as the public and especially the politicians understand the science behind it and sort of what's going on … Information is key, people have to understand what's going on. And both politicians and the general public should understand as much as they can about what the ideas are, what the problems are, what the benefits are, that sort of thing.

Matthew's comments recall the "deficit model" of public communication, where the public's views can be brought in line with those of scientists by providing information and promoting understanding of the science of behavior genetics. He sees his role in preventing eugenic uses of behavior genetics research as one of educating politicians and the public about how genes function and what the potential benefits and problems are with using genes to understand human behavior.

From these brief descriptions of how different scientists describe their stances on public communication, it is clear that there is little consensus in the research community on how researchers should act as managers of the public life of behavior genetics facts. A variety of possible positions that researchers could adopt with respect to public debates about behavior genetics research are all represented here: They could act as advocates in public debates, either by defending the field or by publicly criticizing methods and findings that they believe to be questionable or dangerous; they could act as a neutral mediators who provide education about behavior genetics methods and findings; or researchers could attempt to remove themselves and their findings from public discussion as much as possible.

There is also no clear consensus in the scientific community on what the public believes about genetics or what kinds of facts can travel well in the treacherous terrain of public communication. Many animal behavior geneticists hold the opinion that the public is

essentialist or reductionist in the way that they think about genetics; or as Dennis joked, that the public is "hard wired to believe in genes." These views are aligned with several studies of the presentation of behavior genetics research in the media, which also suggest that news coverage overemphasizes the role of genes. In their study of media coverage on the genetics of alcoholism, for example, Conrad and Weinberg (1996) make the tongue-in-cheek observation that the "gene for" alcoholism has been discovered three times. They demonstrate how news reports announced the discovery of such a gene at three widely separated time points, and argue that in each case the articles overemphasized the contribution of the genes investigated and presented undue optimism about the possibility for cures based on the findings. More recently, Horwitz (2005) has argued that the press coverage of the widely cited Caspi et al. (2003) study framed mental illness as genetic, even though the study demonstrated that the increased risk for depression that they found was the result of the interaction between a particular gene variant and stressful events in early life. But in other cases, researchers described instances where non-scientists were either unaware of genetic contributors to behavioral disorders or actively denied a role for genes in these disorders. Overall, the animal behavior geneticists that I interacted with pointed to relatively few representations of their research and their field that they were satisfied with. In most cases, they described popular representations as too simplistic in either their acceptance or rejection of the basic premises and findings of behavior genetics research.

Hilgartner (1990) argues the dominant view of the popular dissemination of scientific facts characterizes it as a process where scientists create knowledge and then popularizers disseminate simplified accounts of that knowledge. In the best case scenario, the accounts that are circulated in the public sphere are "appropriately" simplified accounts; but at worst, popular accounts of science are seen as "distortions" of genuine scientific knowledge produced by journalists who tend to sensationalize scientific findings or a public that

misunderstands much of what it reads (Hilgartner, 1990). Popularized accounts of science may be produced by journalists or educators, and scientists themselves may also generate simplified statements about their research. When I observed a media training session for Western's drug and alcohol researchers, for example, Eric, the media specialist, encouraged them to "keep it clear, concise, and even simple" when giving an interview, and to "think of some good sound bites ahead of time" that they could convey during interviews. Hilgartner argues that this view of popular dissemination is problematic because scientific experts, as the only actors with access to genuine scientific knowledge, retain the authority to judge what is appropriately simplified and what is distorted, and non-experts "remain forever vulnerable to having their understandings and representations of science derided as 'popularized' and 'distorted'—even if they accurately repeat statements made to them by scientists" (p. 534). In their study of media coverage of genetic research, Bubela and Caulfield (2004) similarly argue that many of the purportedly "hyped" statements about genetics that appear in newspaper articles are actually accurate reflections of claims made in scientific publication and university press releases.

What counts as appropriate simplification and distortion is not always clear, and Hilgartner (1990) argues that observers might make different judgments depending on their social location, interests, or their appraisal of the circumstances. In the Smith laboratory, for example, several members of the lab read genome scientist Craig Venter's autobiographical book, *A Life Decoded* (Venter, 2007), and reactions to Venter's discussion of his family history of alcoholism were mixed. In the book, Venter writes:

> I do enjoy a drink now and then, even though there is a history of alcohol abuse in my family. The complications of alcoholism claimed the life of my grandfather at age sixty-three. His father died while drunk, run over while racing a horse and buggy. Could the susceptibility lie in our dopamine genes? Could my destiny have been shaped by a genetic repetition? In fact, I have four copies of the repeated section of DRD4, which is about average. Other genes are linked with dopamine, so DRD4 does not give the whole picture (p. 31).

Whether or not members of the Smith laboratory judged this portrayal of genetic susceptibility to alcoholism to be "appropriately simple" or "distorted" depended on what message they thought was most important to convey to the public—that alcoholism is a complex disorder, or that alcoholism has a biological basis. Dennis was dissatisfied with this representation of risk for alcoholism because he thought that it did not adequately describe the genetic complexity of addiction. Alcohol drinking does affect dopamine regulation, but he pointed out that it is only one of many brain systems that might be up-regulated or down-regulated in addiction disorders. Marcus, a postdoc in the Smith laboratory, was more sympathetic to Venter's representation of alcoholism. His interpretation was that Venter was overstating the case for DRD4's role in risk for alcoholism in order to get across the point that substance abuse disorders have a biological basis.

I argue that Western researchers' assessment of most popular representations of behavior genetics research as inappropriately simplified in one direction or another suggests that the message that researchers at Western are most concerned with conveying to the public is that behavioral disorders are complex disorders with both environmental and genetic components. Researchers frequently argued that the public's understandings of genetics needed to be complicated because they placed too much emphasis on genes. Hannah, for example, frames her impressions of the public's understanding of behavior genetics by talking about how her family members have overemphasized the importance of heredity in their discussions about a cousin's risk for developing an addiction. She says:

> I have a cousin who was adopted and has a history of drug abuse from his birth parents and all that kind of stuff, and it's like, well, he's clearly going to be susceptible, he's clearly at risk. And it's like yes, but you also don't want to essentially jail someone, lock someone up to try to prevent the inevitable, because you don't know that it's inevitable. So I think people really understand the genetics, but I think they take a very simplistic view of it and don't really realize how flexible, how individual genes have very small effects on it. It's not, you know, you have this gene, you've got it. Upbringing can have a significant impact on that. I think drug experimentation kind of takes on a new stance, too. I'm thinking about this cousin, you know, he had mental health issues,

and it's kind of like well, if he tries something he's going to be an addict.

Hannah describes her extended family's understanding of what it means to be susceptible or at risk for developing a drug abuse problem as too "simplistic" because it ignores the contributions of environmental conditions, such as a positive family upbringing. In her view, the genetic contributors that her cousin may or may not have inherited are more "flexible" and do not inevitably lead to addiction.

In many cases, behavior geneticists expressed the view that their friends, relatives or "the public" held overly determinist views of gene action, but researchers also recalled instances that they portrayed as equally problematic where they interacted with members of the lay public who were "unaware" or actively refused to believe that genetics might contribute to certain behaviors, especially stigmatized disorders such as alcoholism. In one discussion about how researchers incorporate their findings into teaching, I asked Linda, a researcher at a university on the west caoast of the United States, if most of her undergraduate students already held the view that there was a genetic or heritable component to alcoholism. She replied:

> No! Oh my god, I've got this kid who's going to be taking my neuropharm upper levels class, she actually works at a rehab clinic, and she's like I want to do an honors essay, and I said okay, well, pick a neuropsychiatric disease. So she writes me back and she's like, I had no idea, I thought addiction was a disorder of choice. I had no idea that there is a neurobiology underlying it. And I'm like, how did you make it through our biopsych program this far without knowing this? Like I was truly amazed. And glad at the same time, but she says, "I worked at a rehab clinic!" And I'm like, "and it never occurred to you that these people are diseased? It never occurred to you that there might be a neurobiological problem?" So yeah, no, they're not aware.

Her student's substantial exposure to biological conceptions of psychology in her undergraduate courses and in addiction treatment settings made it all the more surprising, in Linda's opinion, that she held the view that addiction was a "disorder of choice." Emily recounts a similar experience from her personal life where she encountered active resistance

to the idea that alcoholism has a genetic basis, despite her attempts to present evidence to the contrary:

> I went on a date like maybe three weeks ago, and I was telling the guy what I did, and he was like, "oh no, alcoholism isn't genetic." And I was like, "but … it is, because I study that! Like for a living." And he's like "no, I don't believe that." And I was like, "well … okay." And I do get a lot of that, or just like, you know, people say because it runs in the family. But when I try to explain, they'll be like, so have you found a gene? Either they don't really believe me or it's so simplified thinking that we're looking for one gene or something, and it's really kind of tough to adequately … adequately explain the complexity of the problem, I guess.

Emily was surprised at her date's negative reaction to her work on the genetics of alcoholism and his refusal to accept her expertise about the disorder. She went on to describe the responses that she often receives from outside the scientific community to behavior genetics research on alcoholism as too genetically deterministic or too focused on free will, but rarely just right: either those that she talks to reject the very idea that behavioral disorders like alcoholism have a genetic basis, or they employ an overly simplistic "one gene, one disorder" model of what the genetics of alcoholism might look like.

While Western researchers may be generally dissatisfied with how members of the public or the popular press represent genetic contributions to behavioral disorders, they also struggle themselves to articulate what kinds of representations should be put in their place. Hannah says that she finds it difficult to say what an "appropriately simplified" message about the genetics of addiction might be:

> We really don't know very much, and so I find it very hard to come up with a definitive answer for myself or to say, even if it's … a simplistic answer. You know, like, well, we think alcohol works through this system. I feel like I need to be like it could also be this, this, this, and this, and this this this, and so what I really need to improve on is kind of the, almost the PR statement. Like, this tells us this. That's not the whole truth, it's not everything, but here's a good statement to take away from this.

Hannah's comments demonstrate Hilgartner's (1990) observation that "genuine" scientific knowledge is often treated as the exclusive preserve of scientists, while non-scientists can

only access simplified representations of that knowledge. Her suggestion that any simplified statement will not be able to convey the "whole truth" about how genes contribute to behavioral disorders makes public communication seem like a somewhat dubious exercise in generating partial truths, and she portrays conveying the complexity of the biology of alcoholism and developing a "PR statement" as mutually exclusive goals. Dennis suggests that the only safe and appropriately simple message to convey to the public is a simplified version of the message of complexity itself. He says:

> The lay public needs to be educated that genes do not determine behavior, and that environments don't determine behavior, but that genes predispose to more or less of a behavior and environment also does the same thing, and the interaction is what determines the behaviors. So yeah, every interview I've ever given with a reporter has probably asked that question, and I've given them that answer. We just have to keep doing it until people understand it.

As I described in chapter 1, this type of claim about the complex nature of behavioral disorders can be thought of as a way of formulating expectations about what kinds of genetic contributions the field is looking for, how quickly they can do so, and what kinds of methods and controls are needed to make them materialize. Delivered in more public settings, the message that both genes and environment contribute to behavior counters both accurately summarizes the current state of knowledge about the genetics of behavioral disorders (at least in Dennis's opinion), and attempts to manage the expectations of imagined publics who might be inclined to place too much importance on either genes or environments.

Finally, in addition to inappropriate simplifications of the content of behavior genetics research, Western scientists also identify inappropriate simplifications of the distance to future applications of behavior genetics research in public discussions about their work. Many researchers argued that the public and the popular press tend not only to simplify the relationship between genes and behaviors, but also to skew discussions towards talking about treatment. As Laura describes it, media interviews tend to jump from background

information about addiction disorders directly to questions about future applications. She says:

> They want to know the basics, you know, how is methamphetamine made? Why is it so addictive? If you do identify genes that are involved, what will that do for us? How do you envision this helping? Are you … they always of course want to know if you're heading towards, is it a treatment so that you can come up with drugs, or is it genetic engineering? Do you plan on changing people's genes in order to help? And so those sorts of questions.

Hannah describes a similar pattern in interactions with her friends and family or in grant writing:

> I think a lot of times especially for addiction, after they get enough over the silly drunk mice thought, which is the first thing, "hahaha, drunk mice" … So after the silly factor, then it's kind of like how does this, you know, help medicine? And so you kind of have to come up with, even if you really don't care, you know, for grant applications or just talking to people you have to kind of come up with a way that this will affect treatment because of this.

Many researchers at Western, like Hannah, are reluctant to present scenarios about how their research might be used in future clinical settings because they see their research as "basic" research that is far removed from these kinds of applications. As Marcus, one of Dennis's postdocs, put it rather overdramatically during the discussion on ethics in the introductory behavior genetics class, he thinks that at present behavior geneticists have no more ability to predict the future of a child born today by looking into his or her genome than a Mayan priest did by divining a child's future by looking into a fire. He argues that researchers have accumulated a lot of descriptive information about what abnormalities look like at the genetic and neural level, but that this information does not yet translate into the ability to predict the course of a disease or to treat it. Marcus's comparison between the present state of behavior genetics research and traditional societies recalls other researchers' descriptions of behavior genetics as a "young" field that will eventually grow up, but is not currently ready to be used for clinical applications.

Eric, the media training expert who does workshops with Western researchers, says that most researchers usually can offer an answer when asked about the relevance of their research, but that relatively few offered this kind of information without being asked directly. George, a researcher at Western who does work both with animal models and human brain imaging, says that in public settings he prefers to talk about the mouse research that he does because people perceive it to be further away from application. When talking about his research on humans with addiction disorders in a public venue, he says that it can get "tricky" quickly because there might be people in the audience with alcoholic relatives or adopted children that have been prenatally exposed to methamphetamine. He says that audience members tend to immediately start relating his research on neuroadaptations in human addicts to their relatives or children, and he worries that this encourages people to adopt a fatalistic stance towards their loved ones' prospects for developing an addiction disorder.

Researchers manage what they perceive to be distortions of the distance to future applications in the behavior genetics field by attempting to expand the field's temporal horizons back out again, pushing clinical scenarios into the far distant future or introducing information about potential complications that makes it seem unlikely that clinical applications will be coming any time soon. Emily, for example, describes a clinical scenario where behavior genetics research could be used to counsel parents about their child's risk factors for behavioral disorders, but she locates such a scenario a hundred years in the future. She says:

> Wouldn't it be cool if in a hundred years, everyone sat down with a genetic counsellor, they did some testing, and they said "your baby is predisposed to X, Y, and Z, here are the environmental things that you could do to protect them against expression of this disease." So you'd get a better shot, I'm not like an advocate of changing the genome, I'm not an advocate of genetic modification in humans, but … you could try to protect people. You could protect people, because they have shown that even if you have genetic predispositions, by being raised in certain environments you will either have less of the disease or it won't show up at all.

This type of scenario, where genetic information is offered to parents in risk counseling sessions, is one that may not necessarily seem so far fetched. Counselors and companies are already using certain types of genetic information to advise those who may be at risk for diseases like breast cancer about their prospects of developing the illness Parthasarathy (2007). But by pushing this scenario as it applies to behavioral disorders a hundred years into the future, Emily moves this scenario from the realm of the plausible into the realm of science fiction.

In addition to displacing future scenarios into the far distant future, researchers also introduced complications into their future scenarios as a way of managing expectations about the clinical applications of their research. Liam, a graduate student at Western working on learning and memory, was one of the few graduate students I interviewed who offered up a scenario describing how his research might be used in the clinic at the beginning of the interview, as part of his description of his work. He says:

> So this has some weight in the clinic because people are really thinking about how to use learning and memory to treat things like post-traumatic stress disorder, or even drug abuse. You can envision someone who's had a traumatic experience coming to the clinic, and while you can't necessarily treat them immediately after they've had the traumatic experience and sort of block some of the emotional effect of that memory, maybe what you can do is ask them, I don't know, give them some sort of script or imagery or something like that that really gets at what their emotional experience was so they're actually retrieving that memory, and then give them some sort of drug like a beta blocker and actually cause the memory to be dampened a little bit so it doesn't have as much of an emotional impact on them.

Liam presents a fairly comprehensive scenario of how his research on learning and memory in mouse models might be used in a clinical setting. Patients who are suffering from post-traumatic stress disorder could be asked to remember the traumatic experience that they endured, and while they are "retrieving" that memory someone could administer a drug that would make that memory more amenable to reformulation so it would not seem as immediate or as traumatic. But he immediately followed this description of a potential

clinical scenario with an explanation of why presenting such a vision was "dangerous." He says:

> And so, I think there's some promise there, maybe. I think it's kind of dangerous because I don't think that it entirely works, and I think that those memories will wind up recurring and it will give some people false hope and things like that … I think that a lot of times folks prematurely jump to the clinic and say that "well this will work in the clinic because of this," those sorts of things. And I think those sorts of assumptions are very dangerous, and that's sort of what got me interested in [memory reconsolidation] is that I think people were sort of abusing the power of the concept and what it could do.

Liam builds up a scenario of future clinical applications only to immediately deconstruct it because he thinks that such scenarios are likely to give the public "false hope" about the promise of behavior genetics research. In the interview, he went on to describe some of the many potential problems he sees with the future scenario that he laid out: that the drug that he is working with would be impractical in a clinical setting because it "smells like poop" and has to be administered intraveneously in large doses; that using drugs to alter a person's emotional responses to a traumatic memory might have side effects on their emotional responses to other things, such as the desire to eat or have sex; and that altering traumatic memories might be taking away experiences that are important for personal growth.

5.4 Conclusion

This chapter explored behavior geneticists' participation in public debates about the meaning and importance of behavior genetics research. At numerous points in its history, the behavior genetics field has become the subject of intense public scrutiny and criticism. In the late 1960s and early 1970s, Arthur Jensen's suggestion that races had different inherited capacities generated extended debates about the potentially racist orientations of the emerging behavior genetics field. Questions about the heritability of intelligence

and differences between human racial groups have continued to re-emerge throughout the field's history, and the publication of Herrnstein and Murray's (1994) *The Bell Curve* and J. Phillipe Rushton's research on the inherited capacities of different races reactivated these debates in the mid-1990s. Other studies, such as Dean Hamer's (1993; 1994) highly publicized findings linking sexual orientation to a region on the X chromosome, have also become the subject of widespread debate by the media, activist groups, and scientists alike.

I argued that these and other past histories have shaped the way that contemporary behavior geneticists frame the public image of their field and imagine the publics that respond to behavior genetics research. Behavior genetics practitioners are engaged in "boundary work" (Gieryn, 1983) on several different fronts, such as the connection of contemporary behavior genetics to past histories of eugenically-oriented research on the heritability of undesirable behavioral traits, the distinction between "scientific" and "ideological" approaches to studying the heritability of behavior, and the connection of current research to potential clinical futures. I also argued that behavior geneticists at Western use these past debates and other public debates about the societal value of scientific research (such as the animal rights movement and debates about global warming) to reason about the political culture in which they work. Using these contentious debates as resources, the publics that researchers at Western imagine are publics that are not only lacking information about how genes and how the scientific method works, but also pontentially capable of distorting and misrepresenting the facts available to them.

Both their participation in crafting the public image of the behavior genetics field and in generating representations of the political culture that they work in contribute to an understanding of the public communication of behavior genetics research as an activity that is difficult and dangerous. I argued that there appears to be no consensus amongst behavior genetics practitioners about how best to manage the contentious terrain of public

communication. While some researchers point to the public or journalists as the source of problematic representations of behavior genetics facts, others feel that perhaps their subject matter is one that is simply inherently controversial. By looking at what researchers at Western identify as "appropriately simplified" and what they see as "distortions" of their work (Hilgartner, 1990), I argued that it is possible to discern a few messages that behavior geneticists at Western feel are safe to impart to the public. These messages are quite similar to some of the main themes of the methodological discussions I have discussed throughout this dissertation about how to conceptualize the relationship between genes and behavior and what kind of information the behavior genetics method can produce. Researchers at Western suggest that what the public (and other researchers) need to understand is that behaviors are complex disorders with multiple environmental and genetic components, neither of which "determine" the behavior in any strong sense. Western researchers also emphasize the long timeline of behavior genetics and push back clinical scenarios, with all of their promise and potential problems, into the far distant future.

This chapter also raises questions for future research about how animal behavior genetics research is portrayed in the media and understood by non-scientists. Some research has explored how laypeople understand human behavior genetics research or how results from human behavior genetics studies are framed in the media (for example, Parens et al., 2006), but little research addresses the public presentation of animal behavior genetics specifically. Is there any difference in the way that animal behavior genetics studies are presented in the media? Do non-scientists actually perceive animal research as further away from clinical applications than human work, as animal behavior geneticists think they do? The many experimental manipulations that can be used to show how genes function in mice might make animal studies more convincing to clinicians or to the public than studies in human populations that show only statistical associations; or conversely, the perceived distance between animals and humans might make animal research seem less valuable.

Additional research on these questions could provide a contrast to the imagined futures and imagined publics of animal behavior geneticists that I have presented here.

Conclusion

Behavior genetics has been a favorite site for discussing the societal implications of contemporary genomics research. The stated goals of behavior genetics—to identify the heritable components of traits like alcoholism, aggression, and intelligence; and even to identify specific genes associated with these behaviors—has inspired many spirited criticisms of the field. The methods and findings of behavior genetics, especially human behavior genetics, have been thoroughly explored and deconstructed by commentators from many disciplines. Stories about the discovery of "genes for" homosexuality, molecular switches that control drinking, and genetic predispositions to aggression have also been popular targets in discussions about the popular dissemination of genetic results and the promises and perils of genetic research.

This dissertation has explored knowledge production in animal behavior genetics from an ethnographic perspective, looking in depth at the day-to-day work of one group of researchers. Throughout this dissertation I have attempted to convey a portrait of what experimental practice is like for researchers in the Smith laboratory: what they worry about in their experimental practice, what they think the process of animal behavior genetics research looks like, and what they hope to eventually produce. Researchers in the Smith laboratory are concerned about managing the "complexity" of behavior through control of the laboratory environment, about the "validity" of their models and their connection to human disorders, about crafting methods that will produce good data, and eventually, about producing information that will translate into improved human

health. They are also concerned that their work might not translate well in particular settings, and that researchers from other disciplines, policy makers, and members of the lay public may misunderstand the subtleties of the findings that they are attempting to produce. Whether or not the Smith laboratory researchers are optimistic about their prospects for accomplishing these goals is difficult to say. On the one hand, experience wih breeding strains of mice that have different tendencies to drink alcohol provides convincing evidence to animal researchers that *something* genetic is going on in drinking behaviors; but on the other hand, the difficult process of trying to identify what that something genetic might be, even in the controlled setting of the laboratory, offers ample evidence that genetic effects on behavior are not straightforward.

Epistemic Scaffolds as a Concept for Understanding Modeling Practices

To help illuminate some of the dynamics of knowledge production in the animal behavior genetics field, in chapter 1 I developed the metaphor of an *epistemic scaffold* to describe the conceptual foundations of particular models or tests. This metaphor draws on research in the history and philosophy of science that describes model organisms as experimentally tractable cases that can be used to understand other organisms or problems, and on research in science and technology studies that describes scientific claims making as a process of generating increasingly more general and risky statements from specific pieces of information. The concept of an "epistemic scaffold" highlights the importance of both of these processes in methodological discussions in animal behavior genetics: Building strong epistemic foundations to support animal behavior genetics research programs involves a horizontal process of linking information from the mouse and the human together, and a vertical process of negotiating the degree of generality and significance of the claims made about the information produced by these models. By incorporating different kinds of

experimental data, theoretical resources, and culturally available information, researchers construct relationships of similarity between mice acting in experimental settings and human behavioral disorders. Widely used tests, such as the elevated plus maze, may have multiple layers of experiments and arguments that add strength and stability to research programs that use these tests as models for particular disorders.

The concept of epistemic scaffolds is useful for understanding several aspects of animal behavior genetics research. First, thinking about epistemic scaffolds helps frame the various methodological arguments, critiques, and suggestions that I documented here as part of larger questions and processes in the field. In chapters 1 and 2, I described some of the aspects of research practice that are important to animal behavior geneticists in the Smith laboratory in their day-to-day work, such as creating a well-controlled laboratory environment and using cautious language to talk about the animal behaviors they are observing in experiments. Complaints about "other" researchers who don't control the lighting in their labs, warnings not to talk in "anthropomorphic" ways, and references to the elevated plus maze as a test for "anxiety-like behavior" can all be seen as ways of acting on the epistemic scaffolds of animal behavior genetics research. Telling cautionary tales about prior research experiences and learning to speak in specific, non-anthropomorphic modes are a few of the ways in which researchers at Western University build local understandings and expectations about the conceptual foundations of their research programs, and indicate to each other that they share particular understandings of what the epistemic underpinnings of animal behavior genetics research should look like.

Researchers can also work on the strength and configuration of epistemic scaffolds in the broader animal behavior genetics community by training practitioners from other laboratories, publishing methodological articles, reviewing each others' articles, and interacting with other researchers around shared models. Some researchers may argue that the elevated plus maze is a test that is still under construction, while others see it

as a test that is ready for widespread use. Some may claim that it is a tool for detecting the effects of anxiolytic drugs, while others use it as a tool for making claims about genes that alter anxiety. These contrasting views and debates indicate differences in the stability that practitioners attribute to particular scaffolds and the traffic practitioners think that these scaffolds were designed to carry. This kind of dynamic in animal behavior genetics knowledge production may be partially visible from methodologically-oriented studies published in the field's literature (such as the study discussed in the introduction that deliberately made the effect of the laboratory environment on different tests apparent through a multi-sited comparison), but ethnographic methods are especially suited to revealing these types of concerns and how practitioners manage them in the interactional setting of the laboratory.

Examining the epistemic scaffolds of particular models also offers a way for analysts to explore how the human is represented in animal behavior genetics research. The representations of the human that are produced through the process of animal modeling are shaped by all of the factors that are highlighted or excluded from the epistemic scaffolding of particular tests. For example, the Alcohol Research Consortium's models for binge drinking represent human drinking as a narrowly defined blood alcohol level that both humans and animals can achieve in a particular period of time. In this model, human behaviors are represented by a quantitative, pharmacological measurement because it is easy to make a cross-species link based on blood alcohol measures, but quite difficult to make links based on other measures (such as the subjective experience of loss of control over one's drinking).

Critics of behavior genetics have argued that this type of research has the tendency to "reduce" humans to genes because of its narrow focus on genetic contributors to behaviors. In chapter 3 I argued that even though animal models are developed for the express purpose of finding genetic contributions to behavioral disorders, researchers themselves

use these models to talk about human behavior in ways that are much more complicated than a simple reductionism would predict. Aspects of the epistemic scaffolds of animal models have metaphorical entailments that are flexible, providing resources for researchers to easily describe environmental factors that structure or motivate human drinking as well as genetic ones. The ways in which the ARC researchers use their binge drinking models to talk about human drinking as environmentally controlled is another example of a process that would be difficult to observe from the published literature alone. Examining how animal behavior geneticists talk about human behaviors in the published literature, where they are aiming to make a case for particular genetic contributions to behavioral disorders, would be unlikely to reveal how researchers use their knowledge from setting up and troubleshooting tests to reflect on other factors shaping human behaviors, and it might present a misleading picture of how much power researchers ascribe to genes.

Finally, the concept of the epistemic scaffold could be useful for understanding modeling practices in other areas of the sciences. Animal behavior genetics is only one of many scientific fields where relationships of similarity are being made and managed; there are many other scientific fields in which researchers are developing "models" of particular phenomena or comparing one type of experimentally tractable object to another object of interest. Developing mathematical models to describe biological processes or studying small groups that are supposed to be representative of larger populations also involves making claims about the similarities of these objects and the utility of particular models as tools for understanding phenomena of interest. Thinking about how researchers construct epistemic scaffolds may also help to illuminate the sociological aspects of modeling in other fields, such as how researchers marshall resources to support their research programs, negotiate the strength of the claims that they make, and selectively represent the phenomena that they are studying.

Knowledge Production under Expectations of Complexity

One of the unique aspects of the case study I have examined is the degree to which researchers expect and emphasize the "complexity" of the phenomena they are studying. In other ethnographic studies of laboratory work, scholars such as Collins (1985) have described knowledge production practices in fields where practitioners are working under the expectation that the natural phenomena that they are studying will present themselves in a regular, law-like fashion. In contrast, the Smith laboratory researchers highlight many instances in which the phenomena they are investigating behave in ways that at first glance appear rather mysterious: tests might generate different results depending on the location where they are conducted or who performs them, researchers produce results that sometimes conflict with prior data even from their own laboratory, results are not considered stable until they have been replicated numerous times and in other tests and laboratories, and findings may or may not translate well to human clincial settings. In this dissertation on knowledge production in animal behavior genetics, I have focused less on how practitioners generate facts from the platforms offered by their epistemic scaffolds, and more on how they manage the excess of uncertainty that they associate with the "complexity" of behavioral disorders.

The assumption that the genetics of human behavioral disorders is likely to be "complex" animates many of the processes that I have described in this dissertation. As I explored in chapter 3, one of the consequences of researchers' understandings of human disorders as "complex" is that they expect that animal models will not be able to capture all of the features of a human behavioral disorder. The strategy that scientists in the alcohol research community have adopted is to develop many animal models for alcoholism, each representing a narrowly-defined feature of the human disorder. Animal behavior geneticists describe reductionism as a methodological strategy that allows them to break down complex disorders for study, rather than a claim about the nature of human behavior.

Researchers also use "complexity" as a way of framing discussions that take place both at Western and in the broader animal behavior genetics field around what the field can reasonably expect to produce, how stable the knowledge that they produce is, and how quickly researchers, funders, or the public can expect knowledge to accumulate. Many of the methodologically-inclined animal behavior geneticists that I interviewed were pushing down on the epistemic scaffolds of animal behavior genetics, trying to encourage their fellow practitioners to adopt what they saw as more defensible and realistic claims in the face of the complex reality of human behavioral disorders. Researchers at Western also expressed concerns that other researchers may be "black boxing" (Latour, 1987) the epistemic scaffolding of behavioral tests, and using them without understanding the specific type of knowledge that the developers of the tests intended them to produce.

The ways in which practitioners modulate the strength of the claims that they make about the relationship of their models to human disorders, particularly in public settings, differentiates this study from other studies of knowledge production processes in science and technology studies. Unlike other studies on the presentation of certainty in the sciences where scholars such as Pinch (1981) have found that scientists present an authoritative air of certainty for public audiences, researchers in the Smith laboratory attempt to circulate their findings with an associated counter-narrative about complexity. In chapter 5 I explored how researchers expressed concerns about the lay public's understanding of behavior genetics results that were similar to the concerns they expressed about practitioners' standards for research practice and assumptions about gene action. While the participants in the discussion and the form of discourse is different in each of these arenas, researchers at Western emphasized similar messages to other researchers, funding agencies, and the public: that behaviors result from complex interactions between genes and environments, and that stable knowledge about the genetics of behavior takes time to produce.

These dynamics of managing complexity could also apply to other scientific fields where

researchers are working under the expectation that the phenomena they are studying are "complex." Nikolas Rose (2007) has argued that a subtle shift towards treating genetics and biology as susceptibilities or probabilistic factors rather than as determinants has taken place in many areas of the biomedical sciences. The processes I have described here of calibrating the strength of claims about research programs, circulating counter-narratives about complexity, and managing expectations about the future timeline of knowledge production may apply to other areas of biomedical research, or even to other scientific fields such as climate sciences where researchers hold similar expectations about the "complexity" of the phenomena that they are studying and are managing similarly contentious intellectual landscapes.

Calling attention to the many actions that researchers in the Smith laboratory take to manage the perceived "complexity" of the phenomena that they are studying also has important implications for social critiques of behavior genetics. As I explored in chapter 3, many critics of genetic research argue that this work is determinist or reductionist in nature, and some critics imply that behavior geneticists themselves hold such ideas about how genes function (see, for example, Duster, 1990; S. Rose, 1997; Rosoff, 2010). Studying animal behavior genetics research ethnographically provides evidence that the environment, far from being absent from researchers' thinking, plays a prominent role in research practice in the Smith laboratory. Researchers at Western are acutely aware of the ways in which the environment can shape behavior, even though their ultimate aim is to create an environment that is controlled enough so that it will be possible to see the effects of particular genes. The narrative of complexity that already exists in the field suggests that critiques of the "reductionist" or "determinist" tendencies of genetic research may no longer be effective interventions in discussions about behavior genetics research. In his work on schizophrenia genetics, Hedgecoe (2001) similarly argues that the "narrative of enlightened geneticization" already incorporates many of the arguments

made by critics of schizophrenia genetics, such as the limitations of current methodologies and the improbability of finding a "gene for" the disorder.

One of the ways in which commentators could intervene in animal behavior genetics research in particular is by elucidating the specific links that are made between the mouse and the human; that is, by participating themselves in configuring animal behavior genetics' epistemic scaffolds. N. Rose and Rached (2009), for example, suggest that animal behavior geneticists could benefit from understanding arguments made in the history and sociology of psychiatry about the classification of human disorders. Healy's (1997) discussion of the pharmaceutical industry's influence in classifying particular drugs and drug effects as "anti-depressant" or "anxiolytic" could be useful for understanding the human categories and drug effects that animal behavior geneticists are linking to in the epistemic scaffolds of tests like the elevated plus maze.

Future Research

As is the case with all ethnographic studies, this study is limited by its focus on a small group of researchers and a particular moment in time in the animal behavior genetics field. The discussion in chapter 1 of the field's changing perceptions of how well-controlled the mouse genome is suggests that expectations about how stabilized aspects of experimental systems are can change within a relatively short period of time. In the mid-1990s, discussions about the genetic background effect portrayed the mouse genome as an object that could generate unexpected results and needed to be further controlled; today, the field's understanding of mouse genetics as well-controlled is again being challenged by researchers pointing to the presence of small but significant duplications in the mouse genome. The field's expectations of how "complex" behavior is and how difficult it will be to produce knowledge about it may also change over time. The researchers that I studied also may not be representative of the views of the animal behavior genetics field as a whole. Since I selected for researchers

who were active in particular methodological discussions in order to get a clearer view of the knowledge production practices of the field, it is likely that my description of experimental life at Western University overstates the extent to which animal behavior geneticists in general pay attention to these methodological issues in their laboratories. Western researchers' descriptions of themselves as particularly "anal" about methods and making links between the animal and the human, especially compared to other practitioners in the field, suggests that the culture of claims making I have described may not be the same elsewhere.

One way this dissertation could be extended is by studying animal behavior genetics researchers from different disciplinary backgrounds, or researchers working in different institutional or national contexts In chapter 4, I described some of the ways in which institutional, commercial, and regulatory factors shape the way that the ideal experimental mouse is conceptualized at Western. Different regulatory systems for managing animal welfare or the introduction of new concepts like "environmental enrichment" into mouse laboratories might subtly shape the epistemic cultures of animal behavior genetics research. Chapter 5 also suggests that researchers in other locations might draw on different histories to construct ideas about the publics they are speaking to and the political culture they are working in.

Future research could also examine how animal behavior genetics knowledge and models travel to new knowledge communities. Epistemic scaffolds are not independent structures that are attached to particular tests; rather, they are generated within particular knowledge communities and their configurations are supported by particular social practices. The cautionary tales that researchers at Western tell and the careful language that they employ when talking about the mouse helps to keep the epistemic scaffolding of particular models open and visible, and to generate shared understandings of the aims of animal behavior genetics research. Western researchers' arguments that practitioners from

other disciplines have a tendency to use behavioral tests in ways that are not "thoughtful" suggests that the epistemic scaffolds of animal behavior genetics models may be becoming increasingly black boxed over time or when tests or findings are moved to new knowledge communities. Future research could explore the extent to which animal behavior genetics research is taken up in different knowledge communities, and investigate how this knowledge changes shape as it moves to different sites.

A Informants and Interviews

The table below provides a complete list of all of the individuals interviewed or identified by name in this project, a general description of their position and location, the date of the recorded interview (if applicable), and the pseudonym that they are identified with in the text of the dissertation (if applicable). Asterisks indicate interviews that were not audio-recorded but documented in field notes afterwards. The names of all field sites, projects, consortiums, and some protocols have also been replaced by generic pseudonyms, such as "Western University" (with the exception of the Jackson Laboratory and the Mouse Phenome Project, which were difficult to anonymize because of the unique features of the site and project).

I chose to use pseudonyms in this project to encourage more candid responses in interviews (especially from graduate students and technicians), and to address concerns raised by some individuals that identifying them by name might make them targets of animal right activism.[1] As is the case for many ethnographies of contemporary science, the convention of using pseudonyms is complicated by the fact that individual scientists are also in some sense public figures whose names appear in publicly available documents such as publications, press releases, or news articles. I chose to use a mixture of real names and pseudonyms to address this problem. In almost all instances, I use pseudonyms to identify individuals when citing from interviews or field notes, even when the individuals being interviewed were willing to be identified by name. Using pseudonyms consistently miti-

[1] See Gusterson, 1996 for a discussion of naming practices in ethnographic studies of controversial scientific fields.

gates against the possibility of identifying individuals who wanted to remain anonymous by association (such as students of a principal investigator who wanted to be interviewed "on the record"). When citing information from scientific publications or other publicly available documents, I use the real names of the authors. This means that in a few cases an individual may be indentified both by his/her real name and by a pseudonym at different places in the text. In some instances, such as my discussion of the research conducted by the "Alcohol Research Consortium," I chose to omit references to the scientific literature entirely to avoid creating compromising links between real names and pseudonyms.

Western University

Dennis Smith	behavior geneticist	August 2006, September 2007, April 2008
Laura Martin	behavior geneticist	March 2008, April 2008
Ruth Tremblay	behavior geneticist	February 2008
Brian McGraw	behavior geneticist	April 2008
Sharon	behavior geneticist	September 2007
Susan	behavioral neuroscientist	May 2008*
	behavioral neuroscientist	May 2008*
Elizabeth	behavioral neuroscientist	May 2008*
George	behavioral neuroscientist	
	technician	September 2007
Alexis	technician	
Delores	technician	
Rachel	technician	
Jack	behavior geneticist	
	lab manager	September 2007
	lab manager	April 2008
Eric	media training specialist	May 2008*
Olivia	information technology manager	September 2007
Aiden	animal care staff	May 2008
Grace	postdoc	September 2007
Marcus	postdoc	
Liam	graduate student	March 2008
	graduate student	March 2008
Sophia	graduate student	March 2008
Emily	graduate student	March 2008
Hannah	graduate student	March 2008
	graduate student	March 2008

Alex	graduate student	April 2008
Chloe	graduate student	April 2008
Matthew	graduate student	April 2008
Ava	graduate student	May 2008

The Jackson Laboratory

Molly Bogue	MPP director	January 2008
Karen	MPP staff	January 2008
Nancy	geneticist	January 2008
	phenotyping center coordinator	January 2008*
	genetic resources manager	January 2008

United States

Frank	behavior geneticist	June 2008
Larry	behavior geneticist	December 2008
	behavioral neuroscientist	December 2008*
Linda	behavior geneticist	November 2009
Charles	behavior geneticist	January 2009
Scott	behavior geneticist	August 2006
Rebecca	behavior geneticist	November 2007
Amy	veterinarian	October 2008
Ethan	graduate student	December 2008
	technician	December 2008
	phenotyping center coordinator	June 2009
	phenotyping center staff	June 2009
	phenotyping center staff	June 2009
	phenotyping center staff	June 2009
Raymond	NIAAA official	June 2009
	NIDA official	June 2009

Canada

Anthony	behavior geneticist	July 2006
William	graduate student	July 2006
David	behavior geneticist	March 2009
Gary	behavior geneticist	January 2009

Germany

Thomas	behavior geneticist	August 2009
	behavior geneticist	August 2009
Mathias	animal behaviorist	August 2007

References

Abbott, A. (1988). *The system of professions: An essay on the division of expert labor*. Chicago, IL: University Of Chicago Press.

Alper, J. S., & Beckwith, J. (1993). Genetic fatalism and social policy: The implications of behavior genetics research. *The Yale Journal of Biology and Medicine*, *66*(6), 511–524.

Ankeny, R. A. (1997). *The conqueror worm: An historical and philosophical examination of the use of the nematode* Caenorhabditis elegans *as a model organism*. Doctoral dissertation, University of Pittsburgh, Graduate Faculty of Arts and Sciences, Pittsburgh, PA.

Ankeny, R. A. (2000). Fashioning descriptive models in biology: Of worms and wiring diagrams. *Philosophy of Science*, *67*, S260–S272.

Ankeny, R. A. (2007). Wormy logic: Model organisms as case-based reasoning. In A. N. H. Creager, L. Elizabeth, & M. N. Wise (Eds.), *Science without laws: Model systems, cases, exemplary narratives* (p. 46-58). Durham, NC: Duke University Press.

Arribas-Ayllon, M., Bartlett, A., & Featherstone, K. (2010). Complexity and accountability: The witches' brew of psychiatric genetics. *Social Studies of Science*, *40*(4), 499–524.

ATLAS.ti Scientific Software Development GmbH. (2010). *ATLAS.ti 6.2*. [Computer software]. Available from `http://www.atlasti.com/`

Bailey, L. (2005). Control and desire: The issue of identity in popular discourses of addiction. *Addiction Research & Theory*, *13*(6), 535–543.

Barnett, J., Burningham, K., Walker, G., & Cass, N. (2010). Imagined publics and engagement around renewable energy technologies in the UK. *Public Understanding of Science*. Available from `http://dx.doi.org/10.1177/0963662510365663` (Published online before print June 30, 2010)

Becker, H. S. (1953). Becoming a marihuana user. *The American Journal of Sociology*, *59*(3), 235–242.

Bennett, E. L., Diamond, M. C., Krech, D., & Rosenzweig, M. R. (1964). Chemical and anatomical plasticity of brain: Changes in brain through experience, demanded by learning theories, are found in experiments with rats. *Science*, *146*(3644), 610–619.

Bogue, M. A., & Grubb, S. C. (2004). The mouse phenome project. *Genetica*, *122*(1), 71–74.

Brannigan, A. (2004). *The rise and fall of social psychology: The use and misuse of the experimental method*. Hawthorne, NY: Aldine de Gruyter.

Bubela, T. M., & Caulfield, T. A. (2004). Do the print media "hype" genetic research? A comparison of newspaper stories and peer-reviewed research papers. *CMAJ*, *170*(9), 1399–1407.

Cambrosio, A., & Keating, P. (1988). "Going monoclonal": Art, science, and magic in the day-to-day use of hybridoma technology. *Social Problems*, *35*(3), 244–260.

Campbell, N. D. (2000). *Using women: Gender, drug policy, and social justice*. New York, NY: Routledge.

Campbell, N. D. (2007). *Discovering addiction: The science and politics of substance abuse research*. Ann Arbor: University of Michigan Press.

Campbell, N. D. (2010). Toward a critical neuroscience of 'addiction'. *BioSocieties*, *5*, 89–104.

Carobrez, A., & Bertoglio, L. (2005). Ethological and temporal analyses of anxiety-like behavior: The elevated plus-maze model 20 years on. *Neuroscience & Biobehavioral*

Reviews, *29*(8), 1193–1205.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford, United Kingdom: Oxford University Press.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, *297*(5582), 851–854.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science*, *301*(5631), 386–389.

Chadarevian, S. de, & Hopwood, N. (Eds.). (2004). *Models: The third dimension of science*. Stanford, CA: Stanford University Press.

Chesler, E. J., Wilson, S. G., Lariviere, W. R., Rodriguez-Zas, S. L., & Mogil, J. S. (2002). Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neuroscience & Biobehavioral Reviews*, *26*(8), 907–923.

Clarke, A. (2005). *Situational analysis: Grounded theory after the postmodern turn*. Thousand Oaks, CA: Sage Publications.

Collins, H. M. (1975). The seven sexes: A study in the sociology of a phenomenon, or the replication of experiments in physics. *Sociology*, *9*(2), 205 –224.

Collins, H. M. (1981). Son of seven sexes: The social destruction of a physical phenomenon. *Social Studies of Science*, *11*(1), 33–62.

Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice*. London, United Kingdom: Sage Publications.

Conrad, P. (1999a). A mirage of genes. *Sociology of Health & Illness*, *21*(2), 228–241.

Conrad, P. (1999b). Uses of expertise: Sources, quotes, and voice in the reporting of genetics in the news. *Public Understanding of Science*, *8*(4), 285–302.

Conrad, P., & Schneider, J. W. (1980). *Deviance and medicalization, from badness to*

sickness. St. Louis, MO: Mosby.

Conrad, P., & Weinberg, D. (1996). Has the gene for alcoholism been discovered three times since 1980? A news media analysis. *Perspectives on Social Problems*, *8*, 3–26.

Cooper, R. M., & Zubek, J. P. (1958). Effects of enriched and restricted early environments on the learning ability of bright and dull rats. *Canadian Journal of Psychology*, *12*(3), 159–164.

Crabbe, J. C., Belknap, J. K., & Buck, K. J. (1994). Genetic animal models of alcohol and drug abuse. *Science*, *264*(5166), 1715–1723.

Crabbe, J. C., & Morris, R. G. M. (2004). Festina lente: Late-night thoughts on high-throughput screening of mouse behavior. *Nature Neuroscience*, *7*(11), 1175–1179.

Crabbe, J. C., Wahlsten, D., & Dudek, B. C. (1999). Genetics of mouse behavior: Interactions with laboratory environment. *Science*, *284*(5420), 1670–1672.

Crawley, J. N. (2007). *What's wrong with my mouse: Behavioral phenotyping of transgenic and knockout mice* (2nd ed.). Hoboken, NJ: Wiley.

Crawley, J. N., & Goodwin, F. K. (1980). Preliminary report of a simple animal behavior model for the anxiolytic effects of benzodiazepines. *Pharmacology Biochemistry and Behavior*, *13*(2), 167–170.

Creager, A. N. H. (2002). *The life of a virus: Tobacco mosaic virus as an experimental model, 1930–1965*. Chicago, IL: University of Chicago Press.

Creager, A. N. H., Lunbeck, E., & Wise, M. N. (2007). *Science without laws: Model systems, cases, exemplary narratives*. Durham, NC: Duke University Press.

Crist, E. (1999). *Images of animals: Anthropomorphism and animal mind*. Philadelphia, PA: Temple University Press.

Cwiartka, M. (forthcoming). The mousification of biomedical research.

Davies, G. (2010). Captivating behaviour: Mouse models, experimental genetics and reductionist returns in the neurosciences. *The Sociological Review*, *58*, 53–72.

Davies, S. R. (2008). Constructing communication. *Science Communication*, *29*(4), 413–434.

Dawson, G. R., Fellow, R., Tricklebank, M. D., & Director, A. (1995). Use of the elevated plus maze in the search for novel anxiolytic agents. *Trends in Pharmacological Sciences*, *16*(2), 33–36.

Degler, C. N. (1991). *In search of human nature: The decline and revival of Darwinism in American social thought*. New York, NY: Oxford University Press.

Despret, V. (2004). The body we care for: Figures of anthropo-zoo-genesis. *Body & Society*, *10*(2-3), 111–134.

Drake, J. W., Charlesworth, B., Charlesworth, D., & Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics*, *148*(4), 1667–1686.

Duster, T. (1990). *Backdoor to eugenics*. New York, NY: Routledge.

Ever-happy mice may hold key to new treatment of depression. (2006, August 8). *Science Daily*. Available from `http://www.sciencedaily.com/releases/2006/08/060822180641.htm`

Fingarette, H. (1988). *Heavy drinking: The myth of alcoholism as a disease*. Berkeley: University of California Press.

Flint, J., Valdar, W., Shifman, S., & Mott, R. (2005). Strategies for mapping and cloning quantitative trait genes in rodents. *Nature Review Genetics*, *6*(4), 271–286.

Frazzetto, G., & Gross, C. (2007). Beyond susceptibility: Behavioural genetics can advance our understanding of psychiatric disorders, but might not meet the expectations for new cures. *EMBO Reports*, *8 Spec No*, S3–6.

Freese, J., & Powell, B. (2003). Tilting at twindmills: Rethinking sociological responses to behavioral genetics. *Journal of Health and Social Behavior*, *44*(2), 130–135.

Frigg, R., & Hartmann, S. (2006). *Models in science.* Stanford Encyclopedia of Philosophy. Available from `http://plato.stanford.edu/entries/models-science/`

Fujimura, J. H. (1987). Constructing 'do-able' problems in cancer research: Articulating alignment. *Social Studies of Science*, *17*(2), 257–293.

Fujimura, J. H. (1996). *Crafting science: A sociohistory of the quest for the genetics of cancer*. Cambridge, MA: Harvard University Press.

Fuller, J. (1960). *Behavior genetics*. New York, NY: Wiley.

Garfinkel, H. (1956). Conditions of successful degradation ceremonies. *The American Journal of Sociology*, *61*(5), 420–424.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Gerlai, R. (1996). Gene-targeting studies of mammalian behavior: is it the mutation or the background genotype? *Trends in Neurosciences*, *19*(5), 177–81.

Gerovitch, S. (2002). *From newspeak to cyberspeak: A history of Soviet cybernetics*. Cambridge, MA: MIT Press.

Gieryn, T. F. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, *48*(6), 781–795.

Gieryn, T. F., & Figert, A. E. (1986). Scientists protect their cognitive authority: The status degredation ceremony of Sir Cyril Burt. In G. Böhme & N. Stehr (Eds.), *The knowledge society: The growing impact of scientific knowledge on social relations* (Vol. 10, pp. 67–86). Norwell, MA: Kluwer Academic Publishers.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine Pub. Co.

Glaser, B. G., & Strauss, A. L. (1968). *Time for dying*. Chicago, IL: Aldine.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. Garden City, NY: Anchor Books.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *The American Journal of Psychiatry*, *160*(4), 636–45.

Gould, S. J. (1981). *The mismeasure of man*. New York, NY: Norton.

Gusfield, J. R. (1981). *The culture of public problems: Drinking-driving and the symbolic order*. Chicago, IL: University of Chicago Press.

Gusfield, J. R. (1986). *Symbolic crusade: Status politics and the American temperance movement* (2nd ed.). Urbana: University of Illinois Press.

Gusterson, H. (1996). *Nuclear rites: A weapons laboratory at the end of the Cold War*. Berkeley: University of California Press.

Gusterson, H. (2008). Nuclear futures: Anticipatory knowledge, expert judgment, and the lack that cannot be filled. *Science and Public Policy*, *35*(8), 551–560.

Hagenbuch, N., Feldon, J., & Yee, B. K. (2006). Use of the elevated plus-maze test with opaque or transparent walls in the detection of mouse strain differences and the anxiolytic effects of diazepam. *Behavioural Pharmacology*, *17*(1), 31–41.

Hamer, D. H. (2002). Genetics: Rethinking behavior genetics. *Science*, *298*(5591), 71–72.

Hamer, D. H., & Copeland, P. (1994). *The science of desire: The search for the gay gene and the biology of behavior*. New York, NY: Simon & Schuster.

Hamer, D. H., Hu, S., Magnuson, V. L., Hu, N., & Pattatucci, A. M. (1993). A linkage between DNA markers on the X chromosome and male sexual orientation. *Science*, *261*(5119), 321–327.

Hay, D. A. (2003). Who should fund and control the direction of human behavior genetics? Review of Nuffield Council on Bioethics 2002 report, Genetics and human behaviour: The ethical context. *Genes, Brain and Behavior*, *2*(6), 321–326.

Healy, D. (1997). *The antidepressant era*. Cambridge, MA: Harvard University Press.

Hedgecoe, A. (2001). Schizophrenia and the narrative of enlightened geneticization.

Social Studies of Science, *31*(6), 875 –911.

Herbert, B. (1994, October 26). In America; throwing a curve. *New York Times*. Available from http://www.nytimes.com/1994/10/26/opinion/in-america-throwing-a-curve.html

Herrnstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York, NY: Free Press.

Herzog, H. A. (1988). The moral status of mice. *American Psychologist*, *43*(6), 473–474.

Hesse, M. B. (1966). *Models and analogies in science*. Notre Dame, IN: University of Notre Dame Press.

Hilgartner, S. (1990). The dominant view of popularization: Conceptual problems, political uses. *Social Studies of Science*, *20*(3), 519–539.

Holden, C. (2008). Parsing the genetics of behavior. *Science*, *322*(5903), 892–895.

Horwitz, A. V. (2005). Media portrayals and health inequalities: A case study of characterizations of gene x environment interactions. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *60 Spec No 2*, 48–52.

Hutchinson, E., Avery, A., & Vandewoude, S. (2005). Environmental enrichment for laboratory rodents. *ILAR Journal / National Research Council, Institute of Laboratory Animal Resources*, *46*(2), 148–161.

Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., et al. (2004). Detection of large-scale variation in the human genome. *Nature Genetics*, *36*(9), 949–951.

Immen, W. (1999, June 4). Laid-back Edmonton mice have scientists puzzled. *Globe and Mail*, pp. A1–A6.

Jackson Laboratory. (2009, September). *Jax® patents unique genetic stability program.* Web page. Retrieved 2010-12-01, from http://jaxmice.jax.org/news/2009/09patented.html

Jasanoff, S. (2004). *States of knowledge: The co-production of science and the social order*. New York, NY: Routledge.

Jellinek, E. M. (1960). *The disease concept of alcoholism*. New Haven, CT: College and University Press.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*(1), 1–123.

Kamin, L. J. (1974). *The science and politics of IQ*. Potomac, MD: L. Erlbaum Associates.

Kay, L. E. (2000). *Who wrote the book of life? A history of the genetic code*. Stanford, CA: Stanford University Press.

Keating, P., & Cambrosio, A. (2003). *Biomedical platforms: Realigning the normal and the pathological in late-twentieth-century medicine*. Cambridge, MA: MIT Press.

Keller, E. F. (2000). Models of and models for: Theory and practice in contemporary biology. *Philosophy of Science*, *67*, S72–S86.

Keller, E. F. (2002). *Making sense of life: Explaining biological development with models, metaphors, and machines*. Cambridge, MA: Harvard University Press.

Keller, E. F. (2008). Nature and the natural. *BioSocieties*, *3*(2), 117–124.

Kendler, K. S. (2005). Psychiatric genetics: A methodologic critique. *American Journal of Psychiatry*, *162*(1), 3–11.

Kevles, D. (1985). *In the name of eugenics: Genetics and the uses of human heredity*. New York, NY: Knopf.

Kevles, D. (1992). *The code of codes: Scientific and social issues in the Human Genome Project*. Cambridge, MA: Harvard University Press.

Kim, Y. (2009). *Handbook of behavior genetics*. New York, NY: Springer.

Kleinman, D. L. (1998). Untangling context: Understanding a university laboratory in the commercial world. *Science, Technology & Human Values*, *23*(3), 285–314.

Kleinman, D. L. (2003). *Impure cultures: University biology and the world of commerce*.

Madison: University of Wisconsin Press.

Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.

Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford, United Kingdom: Pergamon Press.

Kohler, R. F. (1994). *The lords of the fly:* Drosophila *genetics and the experimental life*. Chicago, IL: University of Chicago Press.

Kurt, M., Arik, A. C., & Celik, S. (2000). The effects of sertraline and fluoxetine on anxiety in the elevated plus-maze test in mice. *Journal of Basic and Clinical Physiology and Pharmacology*, *11*(2), 173–180.

Laarakker, M. C., Ohl, F., & Lith, H. A. (2008). Chromosomal assignment of quantitative trait loci influencing modified hole board behavior in laboratory mice using consomic strains, with special reference to anxiety-related behavior and mouse chromosome 19. *Behavior Genetics*, *38*(2), 159–184.

Lakoff, G., & Johnson, M. (1981). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Latour, B. (1986). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Latour, B. (1987). *Science in action*. Cambridge, MA: Harvard University Press.

Lemov, R. (2006). *World as laboratory: Experiments with mice, mazes, and men*. New York, NY: Hill and Wang.

Leonelli, S. (2007). Arabidopsis, the botanical drosophila: From mouse cress to model organism. *Endeavour*, *31*(1), 34–38.

Lerner, R. M. (2006). Another nine-inch nail for behavioral genetics! *Human Development*, *49*(6), 336–342.

Leuenberger, C. (2001). Socialist psychotherapy and its dissidents. *Journal of the History*

of the Behavioral Sciences, *37*(3), 261–273.

Levine, H. G. (1985). The discovery of addiction: Changing conceptions of habitual drunkenness in America. *Journal of Substance Abuse Treatment*, *2*(1), 43–57. (Original work published 1978)

Lewontin, R. C. (1992). *Biology as ideology: The doctrine of DNA*. New York, NY: Harper Perennial.

Lewontin, R. C. (2006). The analysis of variance and the analysis of causes. *International Journal of Epidemiology*, *35*(3), 520–525. (Original work published 1974)

Lewontin, R. C., Rose, S., & Kamin, L. J. (1984). *Not in our genes: Biology, ideology, and human nature*. New York, NY: Pantheon Books.

Li, T., Lumeng, L., & Doolittle, D. P. (1993). Selective breeding for alcohol preference and associated responses. *Behavior Genetics*, *23*(2), 163–170.

Lippman, A. (1991, March). Prenatal genetic testing and screening: Constructing needs and reinforcing inequities. *American Journal of Law & Medicine*, *17*(1/2), 15.

Lister, R. G. (1987). The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology*, *92*(2), 180–185.

Lucki, I. (1997). The forced swimming test as a model for core and component behavioral effects of antidepressant drugs. *Behavioural Pharmacology*, *8*(6-7), 523–532.

Lynch, M. E. (1985). *Art and artifact in laboratory science: A study of shop work and shop talk in a research laboratory*. London, United Kingdom: Routledge & Kegan Paul.

Lynch, M. E. (1988). Sacrifice and the transformation of the animal body into a scientific object: Laboratory culture and ritual practice in the neurosciences. *Social Studies of Science*, *18*(2), 265–289.

Lynch, M. E., Cole, S. A., & McNally, R. (2008). *Truth machine: the contentious history of DNA fingerprinting*. Chicago, IL: University of Chicago Press.

MacAndrew, C., & Edgerton, R. B. (1969). *Drunken comportment: A social explanation*.

Chicago, IL: Aldine.

Maranta, A., Guggenheim, M., Gisler, P., & Pohl, C. (2003). The reality of experts and the imagined lay person. *Acta Sociologica*, *46*(2), 150 –165.

Marashi, V., Barnekow, A., Ossendorf, E., & Sachser, N. (2003). Effects of different forms of environmental enrichment on behavioral, endocrinological, and immunological parameters in male mice. *Hormones and Behavior*, *43*(2), 281–292.

McGuffin, P. (2001). Genomics and behavior: Toward behavioral genomics. *Science*, *291*(5507), 1232–1249.

Michael, M., & Birke, L. (1994). Enrolling the core set: The case of the animal experimentation controversy. *Social Studies of Science*, *24*(1), 81 –95.

Michael, M., & Brown, N. (2000). From the representation of publics to the performance of 'lay political science'. *Social Epistemology: A Journal of Knowledge, Culture and Policy*, *14*(1), 3.

Milner, L. C., & Crabbe, J. C. (2008). Three murine anxiety models: Results from multiple inbred strain comparisons. *Genes, Brain and Behavior*, *7*(4), 496–505.

Mody, C. C. M., & Lynch, M. E. (2010). Test objects and other epistemic things: A history of a nanoscale object. *The British Journal for the History of Science*, *43*(03), 423–458.

Moffitt, T. E. (2005). The new look of behavioral genetics in developmental psychopathology: Gene-environment interplay in antisocial behaviors. *Psychological Bulletin*, *131*(4), 533–554.

Mol, A. (2002). *The body multiple: Ontology in medical practice*. Durham, NC: Duke University Press.

Moore, A. (2000). The key is in the genes…? *EMBO reports*, *1*(2), 100–102.

Morgan, M. S., & Morrison, M. (Eds.). (1999). *Models as mediators: Perspectives on natural and social science*. Cambridge, United Kingdom: Cambridge University Press.

Morse, R. M., & Flavin, D. K. (1992). The definition of alcoholism. *JAMA*, *268*(8),

1012–1014.

Moser, P. C. (1989). An evaluation of the elevated plus-maze test using the novel anxiolytic buspirone. *Psychopharmacology*, *99*(1), 48–53.

Moy, S. S., Nadler, J. J., Perez, A., Barbaro, R. P., Johns, J. M., Magnuson, T. R., et al. (2004). Sociability and preference for social novelty in five inbred strains: An approach to assess autistic-like behavior in mice. *Genes, Brain and Behavior*, *3*(5), 287–302.

Murray, F. (forthcoming). The oncomouse that roared: Hybrid exchange strategies as a source of productive tension at the boundary of overlapping institutions. *American Journal of Sociology*.

National Institute on Alcohol Abuse and Alcoholism. (1992). *The genetics of alcoholism*. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism. (Alcohol Alert 18)

National Institute on Alcohol Abuse and Alcoholism. (2003). *The genetics of alcoholism*. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism. (Alcohol Alert 60)

National Institute on Alcohol Abuse and Alcoholism. (2004, Winter). NIAAA Council approves definition of binge drinking. *NIAAA Newsletter*(3), 3. Available from `http://pubs.niaaa.nih.gov/publications/Newsletter/winter2004/Newsletter_Number3.pdf`

Nelkin, D., & Lindee, M. S. (1995). *The DNA mystique: The gene as a cultural icon*. New York, NY: Freeman.

Nordgren, A. (2003). Metaphors in behavioral genetics. *Theoretical Medicine and Bioethics*, *24*(1), 59–77.

Nuffield Council on Bioethics. (2002). *Genetics and human behaviour: The ethical context*. London, United Kingdom: Nuffield Council on Bioethics.

Office of Juvenile Justice and Delinquency Prevention. (2005). *Drinking in America: Myths,*

realities, and prevention policy. Washington, DC: US Department of Justice, Office of Justice Programs Office of Juvenile Justice and Delinquency Prevention. Available from http://www.udetc.org/documents/Drinking_in_America.pdf

Paigen, K., & Eppig, J. T. (2000). A mouse phenome project. *Mammalian Genome*, *11*(9), 715–717.

Panlab. (2010). *Integrated solutions – anxiety.* Web page. Retrieved 2010-12-01, from http://www.panlab.com/panlabWeb/Solution/php/displaySol.php?nameSolution=ANXIETY#contenido1

Panofsky, A. L. (2006). *Fielding controversy : the genesis and structure of behavior genetics*. Doctoral dissertation, New York University, Graduate School of Arts and Science, New York, NY.

Parens, E., Chapman, A. R., & Press, N. (Eds.). (2006). *Wrestling with behavioral genetics: Science, ethics, and public conversation*. Baltimore, MD: Johns Hopkins University Press.

Parthasarathy, S. (2007). *Building genetic medicine: Breast cancer, technology, and the comparative politics of health care*. Cambridge, MA: MIT Press.

Paul, D. (1991). The Rockefeller Foundation and the origins of behavior genetics. In K. R. Benson, J. Maienschein, & R. Rainger (Eds.), *The expansion of American biology* (pp. 262–283). New Brunswick, NJ: Rutgers University Press.

Pellow, S., & File, S. E. (1986). Anxiolytic and anxiogenic drug effects on exploratory activity in an elevated plus-maze: A novel test of anxiety in the rat. *Pharmacology Biochemistry and Behavior*, *24*(3), 525–529.

Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. Chicago, IL: University of Chicago Press.

Pinch, T. J. (1981). The Sun-Set: the presentation of certainty in scientific life. *Social Studies of Science*, *11*(1), 131 –158.

Pinch, T. J. (1985). Towards an analysis of scientific observation: The externality and evidential significance of observational reports in physics. *Social Studies of Science*, *15*(1), 3–36.

Pálsson, G. (2007). *Anthropology and the new genetics*. Cambridge, United Kingdom: Cambridge University Press.

Porsolt, R. D., Bertin, A., & Jalfre, M. (1977). Behavioral despair in mice: A primary screening test for antidepressants. *Archives Internationales de Pharmacodynamie et de Thérapie*, *229*(2), 327–336.

Powell, C., & Miyakawa, T. (2006). Schizophrenia-relevant behavioral testing in rodent models: A uniquely human disorder? *Biological Psychiatry*, *59*(12), 1198–1207.

Rader, K. A. (1998). "The mouse people": Murine genetics work at the Bussey Institution, 1909–1936. *Journal of the History of Biology*, *31*(3), 327–354.

Rader, K. A. (2004). *Making mice: Standardizing animals for American biomedical research, 1900-1955*. Princeton, NJ: Princeton University Press.

Reardon, J. (2001). The human genome diversity project: A case study in coproduction. *Social Studies of Science*, *31*(3), 357–388.

Reardon, J. (2005). *Race to the finish*. Princeton, NJ: Princeton University Press.

Reinarman, C. (2005). Addiction as accomplishment: The discursive construction of disease. *Addiction Research & Theory*, *13*(4), 307–320.

Reuters News. (1999, June 03). Mice study shows genes are not always destiny. *Reuters News*.

Rheinberger, H. (1997). *Toward a history of epistemic things: Synthesizing proteins in the test tube*. Stanford, CA: Stanford University Press.

Room, R. (1974). Governing images and the prevention of alcohol problems. *Preventive Medicine*, *3*(1), 11–23.

Room, R. (1983). Sociological aspects of the disease concept of alcoholism. *Research*

Advances in Alcohol and Drug Problems, *7*, 47–91.

Rose, N. (2007). *The politics of life itself: Biomedicine, power, and subjectivity in the twenty-first century*. Princeton, NJ: Princeton University Press.

Rose, N., & Rached, J. A. (2009, November 14). *What's wrong with your mice?* Lecture at the conference "Translating Behavior: Bridging Clinial and Animal Model Research". Heidelberg, Germany.

Rose, S. (1995). The rise of neurogenetic determinism. *Nature*, *373*(6513), 380–382.

Rose, S. (1997). *Lifelines: Biology beyond determinism*. Oxford, United Kingdom: Oxford University Press.

Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York, NY: Appleton-Century-Crofts.

Rosoff, P. M. (2010). In search of the mommy gene: Truth and consequences in behavioral genetics. *Science, Technology & Human Values*, *35*(2), 200–243.

Rozmiarek, H. (2007). Origins of the IACUC. In J. Silverman, M. A. Suckow, & S. Murthy (Eds.), *The IACUC handbook* (pp. 1–10). Boca Raton, FL: CRC Press.

Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective*. New Brunswick, NJ: Transaction Publishers.

Sapolsky, R. M. (2005). *Monkeyluv: And other essays on our lives as animals*. New York, NY: Scribner.

Sarkar, S. (1998). *Genetics and reductionism*. Cambridge, United Kingdom: Cambridge University Press.

Schaffner, K. F. (2001). Extrapolation from animal models: Social life, sex and super models. In *Theory and method in the neurosciences* (pp. 200–230). Pittsburgh, PA: University of Pittsburgh Press.

Schneider, J. W. (1978). Deviant drinking as disease: Alcoholism as a social accomplishment. *Social Problems*, *25*(4), 361–372.

Schneider, T., Turczak, J., & Przewłocki, R. (2006). Environmental enrichment reverses behavioral alterations in rats prenatally exposed to valproic acid: Issues for a therapeutic approach in autism. *Neuropsychopharmacology*, *31*(1), 36–46.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., et al. (2004). Large-scale copy number polymorphism in the human genome. *Science*, *305*(5683), 525–528.

Shapin, S., & Schaffer, S. (1985). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life*. Princeton, NJ: Princeton University Press.

She, X., Cheng, Z., Zöllner, S., Church, D. M., & Eichler, E. E. (2008). Mouse segmental duplication and copy number variation. *Nature Genetics*, *40*(7), 909–914.

Shmelkov, S. V., Hormigo, A., Jing, D., Proenca, C. C., Bath, K. G., Milde, T., et al. (2010). Slitrk5 deficiency impairs corticostriatal circuitry and leads to obsessive-compulsive-like behaviors in mice. *Nature Medicine*, *16*(5), 598–602.

Shostak, S. (2007). Translating at work: Genetically modified mouse models and molecularization in the environmental health sciences. *Science, Technology & Human Values*, *32*(3), 315–338.

Silva, A., Paylor, R., Wehner, J., & Tonegawa, S. (1992). Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. *Science*, *257*(5067), 206–211.

Silva, M. T., Alves, C. R., & Santarem, E. M. (1999). Anxiogenic-like effect of acute and chronic fluoxetine on rats tested on the elevated plus-maze. *Brazilian Journal of Medical and Biological Research*, *32*(3), 333–339.

Spires, T. L., & Hannan, A. J. (2005). Nature, nurture and neurology: Gene-environment interactions in neurodegenerative disease. *The FEBS Journal*, *272*(10), 2347–2361.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, *19*(3), 387–420.

Stilgoe, J. (2007). The (co-)production of public uncertainty: UK scientific advice on

mobile phone health risks. *Public Understanding of Science*, *16*(1), 45–61.

Szasz, T. S. (1961). *The myth of mental illness: Foundations of a theory of personal conduct*. New York, NY: Hoeber-Harper.

Traweek, S. (1988). *Beamtimes and lifetimes: the world of high energy physicists*. Cambridge, MA: Harvard University Press.

Tucker, W. H. (1994). *The science and politics of racial research*. Urbana: University of Illinois Press.

Tuunainen, J. (2001). Constructing objects and transforming experimental systems. *Perspectives on Science*, *9*(1), 78–105.

Valverde, M. (1998). *Diseases of the will: Alcohol and the dilemmas of freedom*. Cambridge, United Kingdom: Cambridge University Press.

Venter, J. C. (2007). *A life decoded: My genome, my life*. New York, NY: Viking.

Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity-environment interaction. *Behavioral and Brain Sciences*, *13*, 109–161.

Wahlsten, D. (1994). The intelligence of heritability. *Canadian Psychology / Psychologie Canadienne*, *35*(3), 244–260.

Wahlsten, D. (2003). Airbrushing heritability. *Genes, Brain and Behavior*, *2*(6), 327–329.

Wahlsten, D., Metten, P., Phillips, T. J., II, S. L. B., Burkhart-Kasch, S., Dorow, J., et al. (2003). Different data from different labs: Lessons from studies of gene-environment interaction. *Journal of Neurobiology*, *54*(1), 283–311.

Walf, A. A., & Frye, C. A. (2009). Using the elevated plus maze as a bioassay to assess the effects of naturally occurring and exogenously administered compounds to influence anxiety-related behaviors of mice. *Neuromethods*, *42*, 225–246.

Wall, P. M., & Messier, C. (2000). Ethological confirmatory factor analysis of anxiety-like behaviour in the murine elevated plus-maze. *Behavioural Brain Research*, *114*(1–2), 199–212.

Whitney, G. (1995). Ideology and censorship in behavior genetics. *Mankind Quarterly*, *35*(4), 327–342.

Williams, R., Lim, J. E., Harr, B., Wing, C., Walters, R., Distler, M. G., et al. (2009). A common and unstable copy number variant is associated with differences in Glo1 expression and anxiety-like behavior. *PLoS ONE*, *4*(3), e4649.

Windle, M. (2010). A multilevel developmental contextual approach to substance use and addiction. *BioSocieties*, *5*, 124–136.

Würbel, H. (2002). Behavioral phenotyping enhanced: Beyond (environmental) standardization. *Genes, Brain and Behavior*, *1*(1), 3–8.

Würbel, H., & Gardiner, J. (2007). Refinement of rodent research through environmental enrichment and systematic randomization. *NC3Rs*, *9*, 1–9.

Wynne, B. (1991). Knowledges in context. *Science, Technology & Human Values*, *16*(1), 111–121.

Wynne, B. (1992). Uncertainty and environmental learning. *Global Environmental Change*, *2*(2), 111–127.

Yeoman, B. (2003, July). Can we trust research done with lab mice? *Discover*, 64–71.